

## DESCRIPTION

MULTI-CHANNEL ECHO CANCEL METHOD, MULTI-CHANNEL SOUND  
TRANSFER METHOD, STEREO ECHO CANCELLER, STEREO SOUND TRANSFER  
5 APPARATUS AND TRANSFER FUNCTION CALCULATION APPARATUS

### Related Applications

This application is a continuation of PCT  
application No. PCT/JP02/06968, filed July 10, 2002, which is  
10 based upon, and claims priority from, Japanese Patent  
Application No. 2001-211279, filed July 11, 2001 and Japanese  
Patent Application No. 2002-179722, filed June 20, 2002.

### Technical Field

15 [0001]

This invention relates to an echo cancellation  
technique for multi-channel audio signals and a transfer  
function calculation technique, and solves a problem of  
indefinite coefficient by a new technique.

### Background Art

20 [0002]

In two-way stereo audio transmission that is used in  
teleconferencing systems and so forth, the problem of  
25 indefinite coefficient of echo cancellers has conventionally  
been pointed out and, for solving it, there have been

proposed various techniques (see Journal of The Institute of Electronics, Information and Communication Engineers, vol. 81, No. 3, pp. 266-274, March 1998). As one of the techniques for solving the problem of indefinite coefficient, there is a method of reducing the interchannel correlation. As concrete techniques therefor, there have conventionally been proposed the addition of random noise, the correlation removal by filters, the interchannel frequency shift, the use of an interleave comb filter, the nonlinear processing (Laid-Open Patent Publication No. H10-190848) and so forth.

[0003]

According to the foregoing conventional techniques, since original stereo signals are subjected to processing and then reproduced, there has been a problem that deterioration more or less occurs in reproduced signals. Further, when the processing is complicated, a delay occurs in the reproduced signals, so that there has been a problem of difficulty in conversation in the teleconference and so forth. Further, when the processing is complicated and the processing capability of a processing circuit is low, there have been those instances where it is difficult to update a coefficient of an echo canceller in real time while carrying out the echo cancel processing.

This invention provides an echo cancellation technique for multi-channel audio signals and a transfer function calculation technique that have solved the foregoing problems in the conventional techniques.

# Disclosure of the Invention

[0004]

[Inventions of Claims 1 to 8 and Inventions relating  
5 to such Inventions]

A multi-channel echo cancel method of this invention  
is a method wherein, with respect to a space provided therein  
with a plurality of loudspeakers and one or a plurality of  
microphones and forming a plurality of audio transfer systems  
10 in which multi-channel sounds inputted from an outside and  
reproduced by said respective loudspeakers and having a  
correlation with each other are collected by said microphones,  
individual transfer functions of said plurality of audio  
transfer systems or a plurality of composite transfer  
15 functions obtained by suitably combining said individual  
transfer functions are estimated so as to set corresponding  
filter characteristics, respectively, echo cancel signals are  
respectively produced by giving said set filter  
characteristics to corresponding individual signals to be  
20 reproduced by said respective loudspeakers or a plurality of  
composite signals obtained by suitably combining said  
individual signals, and said echo cancel signals are  
subtracted from corresponding individual collected audio  
signals of said one or plurality of microphones, or a  
25 plurality of composite signals obtained by suitably combining  
said individual collected audio signals, thereby performing  
echo cancellation, and wherein, using as reference signals

(representing those signals that are referred to for  
estimating the transfer functions or the composite transfer  
functions) a set of a plurality of low-correlation composite  
signals which correspond to signals obtained by suitably  
5 combining said multi-channel audio signals and which have a  
lower correlation with each other than that between said  
multi-channel audio signals (e.g. suitably combining said  
multi-channel audio signals to produce a plurality of low-  
correlation composite signals having a lower correlation with  
10 each other than that between said multi-channel audio signals  
and using a set of said plurality of low-correlation  
composite signals as reference signals, or directly inputting  
a set of a plurality of low-correlation composite signals  
which correspond to signals obtained by suitably combining  
15 said multi-channel audio signals and which have a lower  
correlation with each other than that between said multi-  
channel audio signals and using the set of said plurality of  
low-correlation composite signals as reference signals, or  
the like), individual transfer functions of the respective  
20 audio transfer systems or a plurality of composite transfer  
functions obtained by suitably combining said individual  
transfer functions are respectively derived, thereby to set  
corresponding filter characteristics. According to this  
invention, using as reference signals a set of a plurality of  
25 low-correlation composite signals which correspond to signals  
obtained by suitably combining multi-channel audio signals  
having a correlation therebetween and which have a lower

correlation with each other than that between such multi-channel audio signals, individual transfer functions of the respective audio transfer systems or a plurality of composite transfer functions obtained by suitably combining such

5 individual transfer functions are respectively derived, and corresponding filter characteristics are set, thereby to enable echo cancellation. In accordance therewith, since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces

10 deterioration, applied to the multi-channel audio signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation

15 of respectively deriving the individual transfer functions of the respective audio transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual transfer functions, using as the reference signals the set of the plurality of low-correlation composite

20 signals, may be, for example, a calculation of respectively deriving the individual transfer functions of the respective audio transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual transfer functions, based on a cross-spectrum calculation

25 between the plurality of low-correlation composite signals and the individual collected audio signals of the microphones, or the plurality of composite signals obtained by suitably

combining said individual collected audio signals. Further,  
the calculation of respectively deriving the individual  
transfer functions of said plurality of audio transfer  
systems or the plurality of composite transfer functions  
5 obtained by suitably combining said individual transfer  
functions, based on said cross-spectrum calculation, may be,  
for example, a calculation of respectively deriving the  
individual transfer functions of said plurality of audio  
transfer systems or the plurality of composite transfer  
10 functions obtained by suitably combining said individual  
transfer functions, by combining said multi-channel audio  
signals through addition or subtraction to produce a  
plurality of low-correlation composite signals having a lower  
correlation with each other than that between said multi-  
15 channel audio signals, deriving cross spectra between said  
plurality of low-correlation composite signals and the  
individual collected audio signals of the microphones, or the  
plurality of composite signals obtained by suitably combining  
said individual collected audio signals, and ensemble-  
20 averaging them in a predetermined time period per cross  
spectrum.

[0005]

A multi-channel echo cancel method of this invention  
is a method wherein, with respect to a space provided therein  
25 with a plurality of loudspeakers and one or a plurality of  
microphones and forming a plurality of audio transfer systems  
in which multi-channel sounds inputted from an outside and

reproduced by said respective loudspeakers and having a correlation with each other are collected by said microphones, individual transfer functions of said plurality of audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions are estimated so as to set corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said set filter characteristics to corresponding individual signals to be reproduced by said respective loudspeakers or a plurality of composite signals obtained by suitably combining said individual signals, and said echo cancel signals are subtracted from corresponding individual collected audio signals of said one or plurality of microphones, or a plurality of composite signals obtained by suitably combining said individual collected audio signals, thereby performing echo cancellation, wherein, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals (e.g. suitably combining said multi-channel audio signals to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals and using a set of said plurality of low-correlation composite signals as reference signals, or directly inputting

a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals and using the set of said plurality of low-correlation composite signals as reference signals, or the like), estimated errors of individual transfer functions of the respective audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions are respectively derived, thereby to update corresponding filter characteristics to values that cancel said estimated errors. According to this invention, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining multi-channel audio signals having a correlation therebetween and which have a lower correlation with each other than that between such multi-channel audio signals, estimated errors of individual transfer functions of the respective audio transfer systems or a plurality of composite transfer functions obtained by suitably combining such individual transfer functions are respectively derived so as to successively update the corresponding filter characteristics to values that cancel such estimated errors, thereby to enable echo cancellation. In accordance therewith, since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the

multi-channel audio signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation  
5 can be conducted. Further, it is possible to update the filter characteristics in real time. The calculation of respectively deriving the estimated errors of the individual transfer functions of the respective audio transfer systems or the plurality of composite transfer functions obtained by  
10 suitably combining said individual transfer functions, using as the reference signals the set of the plurality of low-correlation composite signals, may be, for example, a calculation of respectively deriving the estimated errors of the individual transfer functions of the respective audio  
15 transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual transfer functions, based on a cross-spectrum calculation between said plurality of low-correlation composite signals and echo cancel error signals obtained by subtracting the  
20 echo cancel signals from the corresponding individual collected audio signals of said one or plurality of microphones, or the plurality of composite signals obtained by suitably combining said individual collected audio signals. Further, the calculation of respectively deriving the  
25 estimated errors of the individual transfer functions of said plurality of audio transfer systems or the plurality of composite transfer functions obtained by suitably combining

said individual transfer functions, based on said cross-spectrum calculation, may be, for example, a calculation of respectively deriving the estimated errors of the individual transfer functions of said plurality of audio transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual transfer functions, by combining said multi-channel audio signals through addition or subtraction to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals, deriving cross spectra between said plurality of low-correlation composite signals and the echo cancel error signals obtained by subtracting the echo cancel signals from the corresponding individual collected audio signals of said one or plurality of microphones, or the plurality of composite signals obtained by suitably combining said individual collected audio signals, and ensemble-averaging them in a predetermined time period per cross spectrum. Further, the correlation between said plurality of low-correlation composite signals is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0006]

A multi-channel sound transfer method of this invention is such that, with respect to two spaces each

forming said plurality of audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the multi-channel audio signals, which have been echo-canceled by performing said method, are transmitted  
5 between said two spaces. In accordance therewith, the multi-channel audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be applied to the teleconferencing system or the like.

[0007]

10 A stereo echo cancel method of this invention is a method wherein, with respect to a space provided therein with two loudspeakers and one or two microphones and forming two or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by  
15 said microphones, individual transfer functions of said two or four audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions are estimated so as to set corresponding filter characteristics, respectively, echo  
20 cancel signals are respectively produced by giving said set filter characteristics to corresponding individual signals to be reproduced by said respective loudspeakers or a plurality of composite signals obtained by suitably combining said individual signals, and said echo cancel signals are  
25 subtracted from corresponding individual collected audio signals of said one or two microphones, or a plurality of composite signals obtained by suitably combining said

individual collected audio signals, thereby performing echo  
cancellation, and wherein, using a sum signal and a  
difference signal of said stereo audio signals as reference  
signals, individual transfer functions of said two or four  
5 audio transfer systems or a plurality of composite transfer  
functions obtained by suitably combining said individual  
transfer functions are respectively derived, thereby to set  
corresponding filter characteristics. According to this  
invention, since the sum signal and the difference signal of  
10 the stereo audio signals have a low correlation therebetween,  
the transfer functions of the two or four audio transfer  
systems or their composite transfer functions are  
respectively derived using the sum signal and the difference  
signal as reference signals, so as to set the corresponding  
15 filter characteristics, thereby to enable echo cancellation.  
In accordance therewith, since the stereo signals can be  
reproduced from the loudspeakers with no or less processing,  
which induces deterioration, applied to the stereo signals,  
excellent reproduced tone quality can be achieved. Further,  
20 there is no or only a small delay in reproduced signals.  
Thus, when applying to the teleconferencing system or the  
like, natural conversation can be conducted. The calculation  
of respectively deriving the individual transfer functions of  
said two or four audio transfer systems or the plurality of  
25 composite transfer functions obtained by suitably combining  
said individual transfer functions, using the sum signal and  
the difference signal of said stereo audio signals as the

reference signals, may be, for example, a calculation of  
respectively deriving the individual transfer functions of  
said two or four audio transfer systems or the plurality of  
composite transfer functions obtained by suitably combining  
5 said individual transfer functions, based on a cross-spectrum  
calculation between the sum signal and the difference signal,  
and the individual collected audio signals of the microphones,  
or the plurality of composite signals obtained by suitably  
combining said individual collected audio signals. Further,  
10 the calculation of respectively deriving the individual  
transfer functions of said two or four audio transfer systems  
or the plurality of composite transfer functions obtained by  
suitably combining said individual transfer functions, based  
on said cross-spectrum calculation, may be, for example, a  
15 calculation of respectively deriving the individual transfer  
functions of said two or four audio transfer systems or the  
plurality of composite transfer functions obtained by  
suitably combining said individual transfer functions, by  
deriving cross spectra between the sum signal and the  
20 difference signal of said stereo audio signals and the  
individual collected audio signals of the microphones or the  
plurality of composite signals obtained by suitably combining  
said individual collected audio signals, and ensemble-  
averaging them in a predetermined time period per cross  
25 spectrum.

[0008]

A stereo echo cancel method of this invention is a

method wherein, with respect to a space provided therein with two loudspeakers and one or two microphones and forming two or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by  
5 said microphones, individual transfer functions of said two or four audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions are estimated so as to set corresponding filter characteristics, respectively, echo  
10 cancel signals are respectively produced by giving said set filter characteristics to corresponding individual signals to be reproduced by said respective loudspeakers or a plurality of composite signals obtained by suitably combining said individual signals, and said echo cancel signals are  
15 subtracted from corresponding individual collected audio signals of said one or two microphones, or a plurality of composite signals obtained by suitably combining said individual collected audio signals, thereby performing echo cancellation, and wherein, using a sum signal and a  
20 difference signal of said stereo audio signals as reference signals, estimated errors of individual transfer functions of said two or four audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions are respectively derived,  
25 thereby to update corresponding filter characteristics to values that cancel said estimated errors. According to this invention, since the sum signal and the difference signal of

the stereo audio signals have a low correlation therebetween,  
the estimated errors of the transfer functions of the two or  
four audio transfer systems or their composite transfer  
functions are respectively derived using the sum signal and  
5 the difference signal as reference signals, so as to  
successively update the corresponding filter characteristics  
to the values that cancel the estimated errors, thereby to  
enable echo cancellation. In accordance therewith, since the  
stereo signals can be reproduced from the loudspeakers with  
10 no or less processing, which induces deterioration, applied  
to the stereo signals, excellent reproduced tone quality can  
be achieved. Further, there is no or only a small delay in  
reproduced signals. Thus, when applying to the  
teleconferencing system or the like, natural conversation can  
15 be conducted. The calculation of respectively deriving the  
estimated errors of the individual transfer functions of said  
two or four audio transfer systems or the plurality of  
composite transfer functions obtained by suitably combining  
said individual transfer functions, using the sum signal and  
20 the difference signal of said stereo audio signals as the  
reference signals, may be, for example, a calculation of  
respectively deriving the estimated errors of the individual  
transfer functions of said two or four audio transfer systems  
or the plurality of composite transfer functions obtained by  
25 suitably combining said individual transfer functions, based  
on a cross-spectrum calculation between the sum signal and  
the difference signal of said stereo audio signals and

respective echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the individual collected audio signals of said one or two microphones, or the plurality of composite signals obtained by suitably combining said individual collected audio signals. The calculation of respectively deriving the estimated errors of the individual transfer functions of said two or four audio transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual transfer functions, based on the cross-spectrum calculation between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, may be, for example, a calculation of respectively deriving the estimated errors of the individual transfer functions of said two or four audio transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual transfer functions, by deriving cross spectra between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, and ensemble-averaging them in a predetermined time period per cross spectrum. Further, the correlation between the sum signal and the difference signal of said stereo audio signals is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0009]

A stereo audio transmission method of this invention is such that, with respect to two spaces each forming said four audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the stereo audio signals, which have been echo-canceled by performing said method, are transmitted between said two spaces. In accordance therewith, the stereo audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be applied to the teleconferencing system or the like.

[0010]

In this invention, the plurality of composite signals obtained by suitably combining the individual signals to be reproduced by the respective loudspeakers, and the plurality of low-correlation composite signals used as reference signals {in this specification, "low-correlation composite signals" is used as including the meaning of uncorrelated signals (uncorrelated composite signals)} may be, for example, common signals.

[0011]

[Inventions of Claims 9 to 21 and Inventions relating to such Inventions]

A multi-channel echo cancel method of this invention is a method wherein, with respect to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones and forming a plurality of audio transfer systems in which multi-channel sounds (e.g. multi-channel stereo

sounds such as two-channel, four-channel, ...) inputted from an outside and reproduced by said respective loudspeakers and having a correlation with each other are collected by said microphones, transfer functions of said plurality of audio transfer systems are estimated so as to set corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said filter characteristics to corresponding signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from corresponding collected audio signals of said one or plurality of microphones, thereby performing echo cancellation, and wherein, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals (e.g. suitably combining said multi-channel audio signals to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals and using a set of said plurality of low-correlation composite signals as reference signals, or directly inputting a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals and using the set of said plurality of

low-correlation composite signals as reference signals, or the like), transfer functions of the respective audio transfer systems are respectively derived, thereby to set corresponding filter characteristics. According to this invention, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining multi-channel audio signals having a correlation therebetween and which have a lower correlation with each other than that between such multi-channel audio signals, the transfer functions of the respective audio transfer systems are respectively derived, and corresponding filter characteristics are set, thereby to enable echo cancellation. In accordance therewith, since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the multi-channel audio signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation of respectively deriving the transfer functions of the respective audio transfer systems using as the reference signals the set of the plurality of low-correlation composite signals, may be, for example, a calculation of respectively deriving the transfer functions of the respective audio transfer systems based on a cross-spectrum calculation between the plurality of low-correlation composite signals

and the respective microphone collected audio signals. Further, the calculation of respectively deriving the transfer functions of said plurality of audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of combining said multi-channel audio signals through addition or subtraction to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals, deriving cross spectra between said plurality of low-correlation composite signals and the respective microphone collected audio signals, and ensemble-averaging them in a predetermined time period per cross spectrum to derive a plurality of kinds of composite transfer functions obtained by combining transfer functions of a plurality of suitable systems among said plurality of audio transfer systems, thereby to derive transfer functions of said plurality of audio transfer systems based on said plurality of kinds of composite transfer functions.

[0012]

A multi-channel echo cancel method of this invention is a method wherein, with respect to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones and forming a plurality of audio transfer systems in which multi-channel sounds inputted from an outside and reproduced by said respective loudspeakers and having a correlation with each other are collected by said microphones, transfer functions of said plurality of audio transfer

systems are estimated so as to set corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said filter characteristics to corresponding signals to be reproduced by said respective  
5 loudspeakers, and said echo cancel signals are subtracted from corresponding collected audio signals of said one or plurality of microphones, thereby performing echo cancellation, and wherein, using as reference signals a set of a plurality of low-correlation composite signals which  
10 correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals (e.g. suitably combining said multi-channel audio signals to produce a plurality of low-  
15 correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals and using a set of said plurality of low-correlation composite signals as reference signals, or directly inputting a set of a plurality of low-correlation composite signals  
20 which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals and using the set of said plurality of low-correlation composite signals as reference signals, or  
25 the like), estimated errors of transfer functions of the respective audio transfer systems are respectively derived, thereby to update corresponding filter characteristics to

values that cancel said estimated errors. According to this invention, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining multi-channel audio signals

5 having a correlation therebetween and which have a lower correlation with each other than that between such multi-channel audio signals, the estimated errors of the transfer functions of the respective audio transfer systems are respectively derived so as to successively update the

10 corresponding filter characteristics to the values that cancel said estimated errors, thereby to enable echo cancellation. In accordance therewith, since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration,

15 applied to the multi-channel audio signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. It is also possible to update

20 the filter characteristics in real time. The filter characteristics can be updated, for example, per suitably determined prescribed time period (e.g. time period of performing said ensemble averaging). The calculation of respectively deriving the estimated errors of the transfer

25 functions of the respective audio transfer systems using the set of the plurality of low-correlation composite signals as the reference signals, may be, for example, a calculation of

respectively deriving the estimated errors of the transfer functions of the respective audio transfer systems based on a cross-spectrum calculation between said plurality of low-correlation composite signals and echo cancel error signals  
5 obtained by subtracting the corresponding echo cancel signals from the collected audio signals of said one or plurality of microphones. Further, the calculation of respectively deriving the estimated errors of the transfer functions of said plurality of audio transfer systems based on said cross-  
10 spectrum calculation, may be, for example, a calculation of combining said multi-channel audio signals through addition or subtraction to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals, deriving  
15 cross spectra between said plurality of low-correlation composite signals and the echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the collected audio signals of said one or plurality of microphones, and ensemble-averaging them in a predetermined  
20 time period per cross spectrum to derive a plurality of kinds of transfer function composite estimated errors obtained by combining estimated errors of transfer functions of a plurality of suitable systems among said plurality of audio transfer systems, thereby to derive estimated errors of the  
25 transfer functions of said plurality of audio transfer systems based on said plurality of kinds of transfer function composite estimated errors. Further, the correlation between

said plurality of low-correlation composite signals is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0013]

Further, the calculation of respectively deriving the transfer functions of said plurality of audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of producing a plurality of mutually orthogonal uncorrelated composite signals by applying a principal component analysis to said multi-channel audio signals, deriving cross spectra between said plurality of uncorrelated composite signals and the respective microphone collected audio signals, and ensemble-averaging them in a predetermined time period per cross spectrum, thereby to derive the transfer functions of said plurality of audio transfer systems based on the ensemble-averaged values.

[0014]

Further, the calculation of respectively deriving the estimated errors of the transfer functions of said plurality of audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of producing a plurality of mutually orthogonal uncorrelated composite signals by applying a principal component analysis to said multi-channel audio signals, deriving cross spectra between said plurality of uncorrelated composite signals and

the echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the collected audio signals of said one or plurality of microphones, and ensemble-averaging them in a predetermined time period per cross spectrum, thereby to derive the estimated errors of the transfer functions of said plurality of audio transfer systems based on the ensemble-averaged values. In this case, it may be arranged that double talk in which sounds other than those reproduced by said loudspeakers are inputted into said microphones is detected and, when the double talk is detected, an update period of said filter characteristics is made relatively longer, whereas, when the double talk is not detected, the update period of said filter characteristics is made relatively shorter, so that it is possible to fully converge the estimated errors when the double talk exists, and further, quicken the convergence of the estimated errors when there is no double talk.

[0015]

A multi-channel sound transfer method of this invention is such that, with respect to two spaces each forming said plurality of audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the multi-channel audio signals, which have been echo-canceled by performing said method, are transmitted between said two spaces. In accordance therewith, the multi-channel audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be

applied to the teleconferencing system or the like.

[0016]

A stereo echo cancel method of this invention is a method wherein, with respect to a space provided therein with  
5 two loudspeakers and one or two microphones and forming two or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said microphones, transfer functions of said two or four audio transfer systems are estimated so as to set  
10 corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said filter characteristics to corresponding signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from corresponding collected  
15 audio signals of said one or two microphones, thereby performing echo cancellation, and wherein, using a sum signal and a difference signal of said stereo audio signals as reference signals, transfer functions of said two or four audio transfer systems are respectively derived, thereby to  
20 set corresponding filter characteristics. According to this invention, since the sum signal and the difference signal of the stereo audio signals have a low correlation therebetween, the transfer functions of the two or four audio transfer systems are respectively derived using the sum signal and the  
25 difference signal as reference signals, so as to set the corresponding filter characteristics, thereby to enable echo cancellation. In accordance therewith, since the stereo

signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the stereo signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation of respectively deriving the transfer functions of said two or four audio transfer systems using the sum signal and the difference signal of said stereo audio signals as the reference signals, may be, for example, a calculation of respectively deriving the transfer functions of said two or four audio transfer systems based on a cross-spectrum calculation between the sum signal and the difference signal, and the respective microphone collected audio signals. Further, the calculation of respectively deriving the transfer functions of said two or four audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of deriving cross spectra between the sum signal and the difference signal of said stereo audio signals and the respective microphone collected audio signals, and ensemble-averaging them in a predetermined time period per cross spectrum to derive a plurality of kinds of composite transfer functions obtained by combining transfer functions of a plurality of suitable systems among said two or four audio transfer systems, thereby to derive transfer functions of said two or four audio transfer systems based on said plurality of kinds of composite transfer

functions. Further, the cross-spectrum calculation in case of the four systems may calculate, for example, respective cross spectra between said sum signal and the first microphone collected audio signal, between said sum signal and the second microphone collected audio signal, between said difference signal and the first microphone collected audio signal, and between said difference signal and the second microphone collected audio signal. Further, said composite transfer functions may include, for example, the first composite transfer function that is the sum of a transfer function between the first loudspeaker and the first microphone and a transfer function between the second loudspeaker and the first microphone, the second composite transfer function that is a difference between the transfer function between the first loudspeaker and the first microphone and the transfer function between the second loudspeaker and the first microphone, the third composite transfer function that is the sum of a transfer function between the first loudspeaker and the second microphone and a transfer function between the second loudspeaker and the second microphone, and the fourth composite transfer function that is a difference between the transfer function between the first loudspeaker and the second microphone and the transfer function between the second loudspeaker and the second microphone. The calculation of deriving the transfer functions of said four audio transfer systems may include, for example, a calculation of deriving a transfer function of

the first audio transfer system from the sum of the first composite transfer function and the second composite transfer function, a calculation of deriving a transfer function of the second audio transfer system from a difference between  
5 the first composite transfer function and the second composite transfer function, a calculation of deriving a transfer function of the third audio transfer system from the sum of the third composite transfer function and the fourth composite transfer function, and a calculation of deriving a  
10 transfer function of the fourth audio transfer system from a difference between the third composite transfer function and the fourth composite transfer function.

[0017]

A stereo echo cancel method of this invention is a  
15 method wherein, with respect to a space provided therein with two loudspeakers and one or two microphones and forming two or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said microphones, transfer functions of said two or four  
20 audio transfer systems are estimated so as to set corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said filter characteristics to corresponding signals to be reproduced by said respective loudspeakers, and said echo  
25 cancel signals are subtracted from corresponding collected audio signals of said one or two microphones, thereby performing echo cancellation, and wherein, using a sum signal

and a difference signal of said stereo audio signals as reference signals, estimated errors of transfer functions of said two or four audio transfer systems are respectively derived, thereby to update corresponding filter

5 characteristics to values that cancel said estimated errors.

According to this invention, since the sum signal and the difference signal of the stereo audio signals have a low correlation therebetween, the estimated errors of the

transfer functions of the two or four audio transfer systems

10 are respectively derived using the sum signal and the difference signal as reference signals, so as to successively update the corresponding filter characteristics to the values that cancel the estimated errors, thereby to enable echo

cancellation. In accordance therewith, since the stereo

15 signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the stereo signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the

20 teleconferencing system or the like, natural conversation can be conducted. It is also possible to update the echo cancel coefficients (filter characteristics) in real time. The

filter characteristics can be updated, for example, per suitably determined prescribed time period (e.g. time period  
25 of performing said ensemble averaging). The calculation of respectively deriving the estimated errors of the transfer functions of said two or four audio transfer systems using

the sum signal and the difference signal of said stereo audio signals as the reference signals, may be, for example, a calculation of respectively deriving the estimated errors of the transfer functions of said two or four audio transfer systems based on a cross-spectrum calculation between the sum signal and the difference signal of said stereo audio signals and respective echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the collected audio signals of said one or two microphones.

Further, the calculation of respectively deriving the estimated errors of the transfer functions of said two or four audio transfer systems based on the cross-spectrum calculation between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, may be, for example, a calculation of deriving cross spectra between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, and ensemble-averaging them in a predetermined time period per cross spectrum to derive a plurality of kinds of transfer function composite estimated errors obtained by combining estimated errors of transfer functions of a plurality of suitable systems among said two or four audio transfer systems, thereby to derive estimated errors of the transfer functions of said two or four audio transfer systems based on said plurality of kinds of transfer function composite estimated errors. Further, the cross-spectrum calculation in case of the four systems may calculate, for example,

respective cross spectra between said sum signal and the first echo cancel error signal, between said sum signal and the second echo cancel error signal, between said difference signal and the first echo cancel error signal, and between  
5 said difference signal and the second echo cancel error signal. Further, said transfer function composite estimated errors may include, for example, the first transfer function composite estimated error that is the sum of an estimated error of a transfer function between the first loudspeaker  
10 and the first microphone and an estimated error of a transfer function between the second loudspeaker and the first microphone, the second transfer function composite estimated error that is a difference between the estimated error of the transfer function between the first loudspeaker and the first  
15 microphone and the estimated error of the transfer function between the second loudspeaker and the first microphone, the third transfer function composite estimated error that is the sum of an estimated error of a transfer function between the first loudspeaker and the second microphone and an estimated  
20 error of a transfer function between the second loudspeaker and the second microphone, and the fourth transfer function composite estimated error that is a difference between the estimated error of the transfer function between the first loudspeaker and the second microphone and the estimated error  
25 of the transfer function between the second loudspeaker and the second microphone. The calculation of deriving the estimated errors of the transfer functions of said four audio

transfer systems may include, for example, a calculation of deriving an estimated error of a transfer function of the first audio transfer system from the sum of the first transfer function composite estimated error and the second transfer function composite estimated error, a calculation of  
5 deriving an estimated error of a transfer function of the second audio transfer system from a difference between the first transfer function composite estimated error and the second transfer function composite estimated error, a  
10 calculation of deriving an estimated error of a transfer function of the third audio transfer system from the sum of the third transfer function composite estimated error and the fourth transfer function composite estimated error, and a  
15 calculation of deriving an estimated error of a transfer function of the fourth audio transfer system from a difference between the third transfer function composite estimated error and the fourth transfer function composite estimated error. Further, the correlation between the sum signal and the difference signal of said stereo audio signals  
20 is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0018]

25 A stereo echo cancel method of this invention is a method wherein, with respect to a space provided therein with two loudspeakers and one or two microphones and forming two

or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said microphones, transfer functions of said two or four audio transfer systems are estimated so as to set  
5 corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said filter characteristics to corresponding signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from corresponding collected  
10 audio signals of said one or two microphones, thereby performing echo cancellation, and wherein a principal component analysis is applied to said stereo audio signals to produce two uncorrelated composite signals that are orthogonal to each other, and transfer functions of said two  
15 or four audio transfer systems are respectively derived using a set of said two uncorrelated composite signals as reference signals, thereby to set corresponding filter characteristics. According to this invention, since the mutually orthogonal two signals produced by applying the principal component  
20 analysis to the stereo audio signals are uncorrelated with each other, the transfer functions of said two or four audio transfer systems are respectively derived using such two signals, thereby to set the corresponding filter characteristics to enable echo cancellation. In accordance  
25 therewith, since the stereo signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the stereo signals, excellent

reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation of

5 respectively deriving the transfer functions of said two or four audio transfer systems using the set of said two uncorrelated composite signals as the reference signals may be, for example, a calculation of respectively deriving transfer functions of said two or four audio transfer systems  
10 based on a cross-spectrum calculation between said two uncorrelated composite signals and the respective microphone collected audio signals. Further, the calculation of respectively deriving the transfer functions of said two or four audio transfer systems based on said cross-spectrum  
15 calculation may be, for example, a calculation of deriving cross spectra between said two uncorrelated composite signals and the respective microphone collected audio signals, and ensemble-averaging them in a predetermined time period per cross spectrum derive a plurality of kinds of composite  
20 transfer functions obtained by combining transfer functions of a plurality of suitable systems among said two or four audio transfer systems, thereby to derive transfer functions of said two or four audio transfer systems based on said plurality of kinds of composite transfer functions.

25 [0019]

A stereo echo cancel method of this invention is a method wherein, with respect to a space provided therein with

two loudspeakers and one or two microphones and forming two  
or four audio transfer systems in which stereo sounds  
reproduced by said respective loudspeakers are collected by  
said microphones, transfer functions of said two or four  
5 audio transfer systems are estimated so as to set  
corresponding filter characteristics, respectively, echo  
cancel signals are respectively produced by giving said  
filter characteristics to corresponding signals to be  
reproduced by said respective loudspeakers, and said echo  
10 cancel signals are subtracted from corresponding collected  
audio signals of said one or two microphones, thereby  
performing echo cancellation, and wherein a principal  
component analysis is applied to said stereo audio signals to  
produce two uncorrelated composite signals that are  
15 orthogonal to each other, and estimated errors of transfer  
functions of said two or four audio transfer systems are  
respectively derived using a set of said two uncorrelated  
composite signals as reference signals, thereby to update  
corresponding filter characteristics to values that cancel  
20 said estimated errors. According to this invention, since  
the mutually orthogonal two signals produced by applying the  
principal component analysis to the stereo audio signals are  
uncorrelated with each other, the estimated errors of the  
transfer functions of said two or four audio transfer systems  
25 are respectively derived using such two signals so as to  
successively update the corresponding filter characteristics  
to the values that cancel such estimated errors, thereby to

enable echo cancellation. In accordance therewith, since the stereo signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the stereo signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. It is also possible to update the echo cancel coefficients (filter characteristics) in real time. The filter characteristics can be updated, for example, per suitably determined prescribed time period (e.g. time period of performing said ensemble averaging). The calculation of respectively deriving the estimated errors of the transfer functions of said two or four audio transfer systems using the set of said two uncorrelated composite signals as the reference signals may be, for example, a calculation of respectively deriving estimated errors of transfer functions of said two or four audio transfer systems based on a cross-spectrum calculation between said two uncorrelated composite signals and respective echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the collected audio signals of said one or two microphones. Further, the calculation of respectively deriving the estimated errors of the transfer functions of said two or four audio transfer systems based on the cross-spectrum calculation between said two uncorrelated composite signals and said echo cancel error signals, may be, for example, a

calculation of deriving cross spectra between said two uncorrelated composite signals and said echo cancel error signals, and ensemble-averaging them in a predetermined time period per cross spectrum to derive a plurality of kinds of transfer function composite estimated errors obtained by combining estimated errors of transfer functions of a plurality of suitable systems among said two or four audio transfer systems, thereby to derive estimated errors of the transfer functions of said two or four audio transfer systems based on said plurality of kinds of transfer function composite estimated errors. In this case, it may be arranged that double talk in which sounds other than those reproduced by said loudspeakers are inputted into said microphones is detected and, when the double talk is detected, an update period of said filter characteristics is made relatively longer, whereas, when the double talk is not detected, the update period of said filter characteristics is made relatively shorter, so that it is possible to fully converge the estimated errors when the double talk exists, and further, quicken the convergence of the estimated errors when there is no double talk.

[0020]

A stereo audio transmission method of this invention is such that, with respect to two spaces each forming said four audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the stereo audio signals, which have been echo-canceled by

performing said method, are transmitted between said two spaces. In accordance therewith, the stereo audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be applied to the teleconferencing system or the like.

[0021]

A stereo echo canceller of this invention is a stereo echo canceller wherein, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, an audio signal supplied to the first loudspeaker is subjected to convolution calculations by first and second filter means, respectively, which are provided corresponding to the first and second microphones, so as to produce first and second echo cancel signals, an audio signal supplied to the second loudspeaker is subjected to convolution calculations by third and fourth filter means, respectively, which are provided corresponding to the first and second microphones, so as to produce third and fourth echo cancel signals, echo cancellation is performed by subtracting, using first subtracting means, said first and third echo cancel signals from a collected audio signal of the first microphone, and echo cancellation is performed by subtracting, using second subtracting means, said second and fourth echo cancel signals from a collected audio signal of the second microphone, said stereo echo

canceller comprising: transfer function calculating means for respectively deriving filter characteristics corresponding to transfer functions of said four audio transfer systems based on a cross-spectrum calculation between a sum signal and a  
5 difference signal of stereo audio signals to be reproduced by said respective loudspeakers and the collected audio signals of said respective microphones, thereby to set said derived filter characteristics to corresponding ones of said first to fourth filter means, respectively. It may be arranged, for  
10 example, that the stereo echo canceller of this invention comprises input means for inputting said stereo audio signals; sum/difference signal producing means for producing a sum signal and a difference signal of the stereo audio signals inputted from said input means; and a main signal  
15 transmission system for transmitting the stereo audio signals inputted from said input means to said respective loudspeakers not through said sum/difference signal producing means, wherein said transfer function calculating means derives the filter characteristics corresponding to the  
20 transfer functions of said four audio transfer systems based on the cross-spectrum calculation between the sum signal and the difference signal produced by said sum/difference signal producing means and the respective microphone collected audio signals, and sets the derived filter characteristics to  
25 corresponding ones of said first to fourth filter means, respectively. Alternatively, it may be arranged that the stereo echo canceller comprises input means for inputting

said stereo audio signals; sum/difference signal producing means for producing a sum signal and a difference signal of the stereo audio signals inputted from said input means; and stereo audio signal demodulating means for calculating the sum of and a difference between the sum signal and the difference signal produced by said sum/difference signal producing means so as to recover the original stereo audio signals, wherein the stereo audio signals recovered by said stereo audio signal demodulating means is transmitted to said  
5  
10  
15  
20  
25  
respective loudspeakers, and said transfer function calculating means derives the filter characteristics corresponding to the transfer functions of said four audio transfer systems based on the cross-spectrum calculation between the sum signal and the difference signal produced by said sum/difference signal producing means and the respective microphone collected audio signals, and sets the derived filter characteristics to corresponding ones of said first to fourth filter means, respectively. Alternatively, it may also be arranged that the stereo echo canceller comprises input means for inputting a sum signal and a difference signal of said stereo audio signals; and stereo audio signal demodulating means for calculating the sum of and a difference between said inputted sum signal and difference signal so as to recover the original stereo audio signals, wherein the stereo audio signals recovered by said stereo audio signal demodulating means is transmitted to said respective loudspeakers, and said transfer function

calculating means derives the filter characteristics corresponding to the transfer functions of said four audio transfer systems based on the cross-spectrum calculation between said inputted sum signal and difference signal and  
5 the respective microphone collected audio signals, and sets the derived filter characteristics to corresponding ones of said first to fourth filter means, respectively.

[0022]

A stereo echo canceller of this invention is a  
10 stereo echo canceller wherein, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, an audio signal  
15 supplied to the first loudspeaker is subjected to convolution calculations by first and second filter means, respectively, which are provided corresponding to the first and second microphones, so as to produce first and second echo cancel signals, an audio signal supplied to the second loudspeaker  
20 is subjected to convolution calculations by third and fourth filter means, respectively, which are provided corresponding to the first and second microphones, so as to produce third and fourth echo cancel signals, echo cancellation is performed by subtracting, using first subtracting means, said  
25 first and third echo cancel signals from a collected audio signal of the first microphone, and echo cancellation is performed by subtracting, using second subtracting means,

said second and fourth echo cancel signals from a collected audio signal of the second microphone, said stereo echo canceller comprising: transfer function calculating means for respectively deriving estimated errors of transfer functions of said four audio transfer systems based on a cross-spectrum calculation between a sum signal and a difference signal of stereo audio signals to be reproduced by said respective loudspeakers and respective echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the collected audio signals of said two microphones, thereby to update filter characteristics of said first to fourth filter means to values that cancel said estimated errors, respectively. It may be arranged, for example, that the stereo echo canceller of this invention comprises input means for inputting said stereo audio signals; sum/difference signal producing means for producing a sum signal and a difference signal of the stereo audio signals inputted from said input means; and a main signal transmission system for transmitting the stereo audio signals inputted from said input means to said respective loudspeakers not through said sum/difference signal producing means, wherein said transfer function calculating means derives the estimated errors of the transfer functions of said four audio transfer systems based on the cross-spectrum calculation between the sum signal and the difference signal produced by said sum/difference signal producing means and the respective echo cancel error signals, and updates the filter characteristics

of said first to fourth filter means to the values that cancel said estimated errors, respectively. Alternatively, it may be arranged that the stereo echo canceller comprises input means for inputting said stereo audio signals;

5 sum/difference signal producing means for producing a sum signal and a difference signal of the stereo audio signals inputted from said input means; and stereo audio signal demodulating means for calculating the sum of and a difference between the sum signal and the difference signal

10 produced by said sum/difference signal producing means so as to recover the original stereo audio signals, wherein the stereo audio signals recovered by said stereo audio signal demodulating means is transmitted to said respective loudspeakers, and said transfer function calculating means

15 derives the estimated errors of the transfer functions of said four audio transfer systems based on the cross-spectrum calculation between the sum signal and the difference signal produced by said sum/difference signal producing means and the respective echo cancel error signals, and updates the

20 filter characteristics of said first to fourth filter means to the values that cancel said estimated errors, respectively. Alternatively, it may also be arranged that the stereo echo canceller comprises input means for inputting a sum signal and a difference signal of said stereo audio signals; and

25 stereo audio signal demodulating means for calculating the sum of and a difference between said inputted sum signal and difference signal so as to recover the original stereo audio

signals, wherein the stereo audio signals recovered by said stereo audio signal demodulating means is transmitted to said respective loudspeakers, and said transfer function calculating means derives the estimated errors of the transfer functions of said four audio transfer systems based on the cross-spectrum calculation between said inputted sum signal and difference signal produced by said sum/difference signal producing means and the respective echo cancel error signals, and updates the filter characteristics of said first to fourth filter means to the values that cancel said estimated errors, respectively.

[0023]

A stereo echo canceller of this invention is a stereo echo canceller wherein, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, an audio signal supplied to the first loudspeaker is subjected to convolution calculations by first and second filter means, respectively, which are provided corresponding to the first and second microphones, so as to produce first and second echo cancel signals, an audio signal supplied to the second loudspeaker is subjected to convolution calculations by third and fourth filter means, respectively, which are provided corresponding to the first and second microphones, so as to produce third and fourth echo cancel signals, echo cancellation is

performed by subtracting, using first subtracting means, said first and third echo cancel signals from a collected audio signal of the first microphone, and echo cancellation is performed by subtracting, using second subtracting means, said second and fourth echo cancel signals from a collected audio signal of the second microphone, said stereo echo canceller comprising: transfer function calculating means for respectively deriving filter characteristics corresponding to transfer functions of said four audio transfer systems based on a cross-spectrum calculation between mutually orthogonal two uncorrelated composite signals produced by applying a principal component analysis to stereo audio signals to be reproduced by said respective loudspeakers and the respective microphone collected audio signals, thereby to set said derived filter characteristics to corresponding ones of said first to fourth filter means, respectively. It may be arranged, for example, that the stereo echo canceller of this invention comprises input means for inputting said stereo audio signals; orthogonalizing means for applying a principal component analysis to the stereo audio signals inputted from said input means to produce mutually orthogonal two uncorrelated composite signals; and a main signal transmission system for transmitting the stereo audio signals inputted from said input means to said respective loudspeakers not through said orthogonalizing means, wherein said transfer function calculating means derives the filter characteristics corresponding to the transfer functions of

said four audio transfer systems based on the cross-spectrum calculation between the two uncorrelated composite signals produced by said orthogonalizing means and the respective microphone collected audio signals, and sets the derived  
5 filter characteristics to corresponding ones of said first to fourth filter means, respectively.

[0024]

A stereo echo canceller of this invention is a stereo echo canceller wherein, with respect to a space  
10 provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, an audio signal supplied to the first loudspeaker is subjected to convolution  
15 calculations by first and second filter means, respectively, which are provided corresponding to the first and second microphones, so as to produce first and second echo cancel signals, an audio signal supplied to the second loudspeaker is subjected to convolution calculations by third and fourth  
20 filter means, respectively, which are provided corresponding to the first and second microphones, so as to produce third and fourth echo cancel signals, echo cancellation is performed by subtracting, using first subtracting means, said first and third echo cancel signals from a collected audio  
25 signal of the first microphone, and echo cancellation is performed by subtracting, using second subtracting means, said second and fourth echo cancel signals from a collected

audio signal of the second microphone, said stereo echo  
canceller comprising: transfer function calculating means for  
respectively deriving estimated errors of transfer functions  
of said four audio transfer systems based on a cross-spectrum  
5 calculation between mutually orthogonal two uncorrelated  
composite signals produced by applying a principal component  
analysis to stereo audio signals to be reproduced by said  
respective loudspeakers and respective echo cancel error  
signals obtained by subtracting the corresponding echo cancel  
10 signals from the collected audio signals of said two  
microphones, thereby to update filter characteristics of said  
first to fourth filter means to values that cancel said  
estimated errors, respectively. It may be arranged, for  
example, that the stereo echo canceller of this invention  
15 comprises input means for inputting said stereo audio  
signals; orthogonalizing means for applying a principal  
component analysis to the stereo audio signals inputted from  
said input means to produce mutually orthogonal two  
uncorrelated composite signals; and a main signal  
20 transmission system for transmitting the stereo audio signals  
inputted from said input means to said respective  
loudspeakers not through said orthogonalizing means, wherein  
said transfer function calculating means derives the  
estimated errors of the transfer functions of said four audio  
25 transfer systems based on the cross-spectrum calculation  
between the two uncorrelated composite signals produced by  
said orthogonalizing means and the respective echo cancel

error signals, and updates the filter characteristics of said first to fourth filter means to the values that cancel said estimated errors, respectively. In this case, it may be arranged that double talk detecting means is provided for  
5 detecting double talk in which sounds other than those reproduced by said loudspeakers are inputted into said microphones and, when the double talk is detected, said transfer function calculating means makes relatively longer an update period of said filter characteristics, whereas,  
10 when the double talk is not detected, it makes relatively shorter the update period of said filter characteristics, so that it is possible to fully converge the estimated errors when the double talk exists, and further, quicken the convergence of the estimated errors when there is no double  
15 talk.

[0025]

The stereo echo canceller of this invention may be further provided with correlation detecting means for detecting the correlation between the sum signal and the  
20 difference signal of said stereo audio signals and, when a value of said correlation is no less than a prescribed value, stopping updating of said filter characteristics, thereby to prevent the echo cancel error signals from unexpectedly increasing.

25 [0026]

A stereo sound transfer apparatus of this invention is such that, with respect to two spaces each forming said

four audio transfer systems, any of said stereo echo cancellers is arranged in each space, so that the stereo audio signals, which have been echo-canceled by said stereo echo cancellers, are transmitted between said two spaces.

5 [0027]

[Inventions of Claims 22 to 27 and Inventions relating to such Inventions]

A multi-channel echo cancel method of this invention is a method wherein, with respect to a space provided therein  
10 with a plurality of loudspeakers and one or a plurality of microphones and forming a plurality of audio transfer systems in which multi-channel sounds reproduced by said respective loudspeakers and having a correlation with each other are collected by said microphones, composite transfer functions  
15 of said plurality of audio transfer systems are estimated so as to set corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said set filter characteristics to composite signals of individual signals to be reproduced by said respective loudspeakers, and  
20 said echo cancel signals are subtracted from individual collected audio signals of said one or plurality of microphones, thereby performing echo cancellation, and wherein, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals  
25 obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals (e.g.

suitably combining said multi-channel audio signals to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals and using a set of said  
5 plurality of low-correlation composite signals as reference signals, or directly inputting a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other  
10 than that between said multi-channel audio signals and using the set of said plurality of low-correlation composite signals as reference signals, or the like), composite transfer functions of said plurality of audio transfer systems are respectively derived, thereby to set  
15 corresponding filter characteristics. According to this invention, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining multi-channel audio signals having a correlation therebetween and which have a lower  
20 correlation with each other than that between such multi-channel audio signals, the composite transfer functions of said plurality of audio transfer systems are respectively derived, and corresponding filter characteristics are set, thereby to enable echo cancellation. In accordance therewith,  
25 since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the multi-channel audio signals,

excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation of respectively deriving the composite transfer functions of said plurality of audio transfer systems using as the reference signals the set of the plurality of low-correlation composite signals, may be, for example, a calculation of respectively deriving the composite transfer functions of said plurality of audio transfer systems based on a cross-spectrum calculation between the plurality of low-correlation composite signals and the individual collected audio signals of the respective microphones. Further, the calculation of respectively deriving the composite transfer functions of said plurality of audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of combining said multi-channel audio signals through addition or subtraction to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals, deriving cross spectra between said plurality of low-correlation composite signals and the individual collected audio signals of the respective microphones, and ensemble-averaging them in a predetermined time period per cross spectrum to derive composite transfer functions of said plurality of audio transfer systems.

[0028]

A multi-channel echo cancel method of this invention is a method wherein, with respect to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones and forming a plurality of audio transfer systems in which multi-channel sounds reproduced by said respective loudspeakers and having a correlation with each other are collected by said microphones, composite transfer functions of said plurality of audio transfer systems are estimated so as to set corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said set filter characteristics to composite signals of individual signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from individual collected audio signals of said one or plurality of microphones, thereby performing echo cancellation, and wherein, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals (e.g. suitably combining said multi-channel audio signals to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals and using a set of said plurality of low-correlation composite signals as reference signals, or directly inputting a set of a plurality of low-correlation composite signals which correspond to signals

obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals and using the set of said plurality of low-correlation composite

5 signals as reference signals, or the like), estimated errors of composite transfer functions of said plurality of audio transfer systems are respectively derived, thereby to update corresponding filter characteristics to values that cancel said estimated errors. According to this invention, using as

10 reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining multi-channel audio signals having a correlation therebetween and which have a lower correlation with each other than that between such multi-channel audio

15 signals, the estimated errors of the composite transfer functions of said plurality of audio transfer systems are respectively derived so as to successively update the corresponding filter characteristics to the values that cancel said estimated errors, thereby to enable echo

20 cancellation. In accordance therewith, since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the multi-channel audio signals, excellent

25 reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. It is also possible to update

the filter characteristics in real time. The calculation of respectively deriving the estimated errors of the composite transfer functions of said plurality of audio transfer systems using the set of the plurality of low-correlation

5 composite signals as the reference signals, may be, for example, a calculation of respectively deriving the estimated errors of the composite transfer functions of said plurality of audio transfer systems based on a cross-spectrum calculation between said plurality of low-correlation

10 composite signals and echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the individual collected audio signals of said one or plurality of microphones. Further, the calculation of respectively deriving the estimated errors of the composite transfer

15 functions of said plurality of audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of combining said multi-channel audio signals through addition or subtraction to produce a plurality of low-correlation composite signals having a lower correlation

20 with each other than that between said multi-channel audio signals, deriving cross spectra between said plurality of low-correlation composite signals and the echo cancel error signals obtained by subtracting the corresponding echo cancel

25 signals from the individual collected audio signals of said one or plurality of microphones, and ensemble-averaging them in a predetermined time period per cross spectrum to derive estimated errors of the composite transfer functions of said

plurality of audio transfer systems. Further, the correlation between said plurality of low-correlation composite signals is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0029]

A multi-channel sound transfer method of this invention is such that, with respect to two spaces each forming said plurality of audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the multi-channel audio signals, which have been echo-canceled by performing said method, are transmitted between said two spaces. In accordance therewith, the multi-channel audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be applied to the teleconferencing system or the like.

[0030]

A stereo echo cancel method of this invention is a method wherein, with respect to a space provided therein with two loudspeakers and one or two microphones and forming two or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said microphones, composite transfer functions of said two or four audio transfer systems are estimated so as to set corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said set

filter characteristics to composite signals of individual signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from individual collected audio signals of said one or two microphones,

5 thereby performing echo cancellation, and wherein, using a sum signal and a difference signal of said stereo audio signals as reference signals, composite transfer functions of said two or four audio transfer systems are respectively derived, thereby to set corresponding filter characteristics.

10 According to this invention, since the sum signal and the difference signal of the stereo audio signals have a low correlation therebetween, the composite transfer functions of the two or four audio transfer systems are respectively derived using the sum signal and the difference signal as  
15 reference signals, so as to set the corresponding filter characteristics, thereby to enable echo cancellation. In accordance therewith, since the stereo signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the stereo signals,  
20 excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals.

Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation of respectively deriving the composite transfer functions of  
25 said two or four audio transfer systems using the sum signal and the difference signal of said stereo audio signals as the reference signals, may be, for example, a calculation of

respectively deriving the composite transfer functions of  
said two or four audio transfer systems based on a cross-  
spectrum calculation between the sum signal and the  
difference signal, and the individual collected audio signals  
5 of the respective microphones. Further, the calculation of  
respectively deriving the composite transfer functions of  
said two or four audio transfer systems based on said cross-  
spectrum calculation, may be, for example, a calculation of  
deriving cross spectra between the sum signal and the  
10 difference signal of said stereo audio signals and the  
individual collected audio signals of the respective  
microphones, and ensemble-averaging them in a predetermined  
time period per cross spectrum to derive composite transfer  
functions of said two or four audio transfer systems.

15 [0031]

A stereo echo cancel method of this invention is a  
method wherein, with respect to a space provided therein with  
two loudspeakers and one or two microphones and forming two  
or four audio transfer systems in which stereo sounds  
20 reproduced by said respective loudspeakers are collected by  
said microphones, composite transfer functions of said two or  
four audio transfer systems are estimated so as to set  
corresponding filter characteristics, respectively, echo  
cancel signals are respectively produced by giving said set  
25 filter characteristics to composite signals of individual  
signals to be reproduced by said respective loudspeakers, and  
said echo cancel signals are subtracted from individual

collected audio signals of said one or two microphones,  
thereby performing echo cancellation, and wherein, using a  
sum signal and a difference signal of said stereo audio  
signals as reference signals, estimated errors of composite  
5 transfer functions of said two or four audio transfer systems  
are respectively derived, thereby to update corresponding  
filter characteristics to values that cancel said estimated  
errors. According to this invention, since the sum signal  
and the difference signal of the stereo audio signals have a  
10 low correlation therebetween, the estimated errors of the  
composite transfer functions of the two or four audio  
transfer systems are respectively derived using the sum  
signal and the difference signal as reference signals, so as  
to successively update the corresponding filter  
15 characteristics to the values that cancel the estimated  
errors, thereby to enable echo cancellation. In accordance  
therewith, since the stereo signals can be reproduced from  
the loudspeakers with no or less processing, which induces  
deterioration, applied to the stereo signals, excellent  
20 reproduced tone quality can be achieved. Further, there is  
no or only a small delay in reproduced signals. Thus, when  
applying to the teleconferencing system or the like, natural  
conversation can be conducted. The calculation of  
respectively deriving the estimated errors of the composite  
25 transfer functions of said two or four audio transfer systems  
using the sum signal and the difference signal of said stereo  
audio signals as the reference signals, may be, for example,

a calculation of respectively deriving the estimated errors of the composite transfer functions of said two or four audio transfer systems based on a cross-spectrum calculation between the sum signal and the difference signal of said stereo audio signals and respective echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the individual collected audio signals of said one or two microphones. Further, the calculation of respectively deriving the estimated errors of the composite transfer functions of said two or four audio transfer systems based on the cross-spectrum calculation between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, may be, for example, a calculation of deriving cross spectra between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, and ensemble-averaging them in a predetermined time period per cross spectrum to derive estimated errors of the composite transfer functions of said two or four audio transfer systems. Further, the correlation between the sum signal and the difference signal of said stereo audio signals is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0032]

A stereo audio transmission method of this invention is such that, with respect to two spaces each forming said

four audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the stereo audio signals, which have been echo-canceled by performing said method, are transmitted between said two

5 spaces. In accordance therewith, the stereo audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be applied to the teleconferencing system or the like.

[0033]

10 A stereo echo canceller of this invention is a stereo echo canceller wherein, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are  
15 collected by said respective microphones, a sum signal of stereo audio signals to be reproduced by said respective loudspeakers is subjected to convolution calculations by first and second filter means, respectively, so as to produce first and second echo cancel signals, a difference signal of  
20 the stereo audio signals to be reproduced by said respective loudspeakers is subjected to convolution calculations by third and fourth filter means, respectively, so as to produce third and fourth echo cancel signals, echo cancellation is performed by subtracting, using first subtracting means, said  
25 first and third echo cancel signals from a collected audio signal of the first microphone, and echo cancellation is performed by subtracting, using second subtracting means,

said second and fourth echo cancel signals from a collected audio signal of the second microphone, said stereo echo canceller comprising: transfer function calculating means for respectively deriving filter characteristics corresponding to  
5 composite transfer functions of said four audio transfer systems based on a cross-spectrum calculation between the sum signal and the difference signal of the stereo audio signals to be reproduced by said respective loudspeakers and the respective microphone collected audio signals, thereby to set  
10 said derived filter characteristics to corresponding ones of said first to fourth filter means, respectively. It may be arranged, for example, that the stereo echo canceller of this invention comprises input means for inputting said stereo audio signals; sum/difference signal producing means for  
15 producing a sum signal and a difference signal of the stereo audio signals inputted from said input means; and a main signal transmission system for transmitting the stereo audio signals inputted from said input means to said respective loudspeakers not through said sum/difference signal producing  
20 means, wherein said transfer function calculating means derives the filter characteristics corresponding to the composite transfer functions of said four audio transfer systems based on the cross-spectrum calculation between the sum signal and the difference signal produced by said  
25 sum/difference signal producing means and the respective microphone collected audio signals, and sets the derived filter characteristics to corresponding ones of said first to

fourth filter means, respectively.

[0034]

A stereo echo canceller of this invention is a stereo echo canceller wherein, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, a sum signal of stereo audio signals to be reproduced by said respective loudspeakers is subjected to convolution calculations by first and second filter means, respectively, so as to produce first and second echo cancel signals, a difference signal of the stereo audio signals to be reproduced by said respective loudspeakers is subjected to convolution calculations by third and fourth filter means, respectively, so as to produce third and fourth echo cancel signals, echo cancellation is performed by subtracting, using first subtracting means, said first and third echo cancel signals from a collected audio signal of the first microphone, and echo cancellation is performed by subtracting, using second subtracting means, said second and fourth echo cancel signals from a collected audio signal of the second microphone, said stereo echo canceller comprising: transfer function calculating means for respectively deriving estimated errors of composite transfer functions of said four audio transfer systems based on a cross-spectrum calculation between the sum signal and the difference signal of the stereo audio signals to be

reproduced by said respective loudspeakers and respective echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the collected audio signals of said two microphones, thereby to update filter characteristics of said first to fourth filter means to values that cancel said estimated errors, respectively. It may be arranged, for example, that the stereo echo canceller of this invention comprises input means for inputting said stereo audio signals; sum/difference signal producing means for producing a sum signal and a difference signal of the stereo audio signals inputted from said input means; and a main signal transmission system for transmitting the stereo audio signals inputted from said input means to said respective loudspeakers not through said sum/difference signal producing means, wherein said transfer function calculating means derives the estimated errors of the composite transfer functions of said four audio transfer systems based on the cross-spectrum calculation between the sum signal and the difference signal produced by said sum/difference signal producing means and the respective echo cancel error signals, and updates the filter characteristics of said first to fourth filter means to the values that cancel said estimated errors, respectively.

[0035]

The stereo echo canceller of this invention may be further provided with correlation detecting means for detecting the correlation between the sum signal and the

difference signal of said stereo audio signals and, when a value of said correlation is no less than a prescribed value, stopping updating of said filter characteristics, thereby to prevent the echo cancel error signals from unexpectedly  
5 increasing.

[0036]

A stereo sound transfer apparatus of this invention is such that, with respect to two spaces each forming said four audio transfer systems, any of said stereo echo  
10 cancellers is arranged in each space, so that the stereo audio signals, which have been echo-canceled by said stereo echo cancellers, are transmitted between said two spaces.

[0037]

[Inventions of Claims 28 to 31 and Inventions  
15 relating to such Inventions]

A multi-channel echo cancel method of this invention is a method wherein, with respect to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones and forming a plurality of audio transfer systems  
20 in which multi-channel sounds reproduced by said respective loudspeakers and having a correlation with each other are collected by said microphones, composite transfer functions of said plurality of audio transfer systems are estimated so as to set corresponding filter characteristics, respectively,  
25 echo cancel signals are respectively produced by giving said set filter characteristics to individual signals to be reproduced by said respective loudspeakers, and said echo

cancel signals are subtracted from composite signals of individual collected audio signals of said one or plurality of microphones, thereby performing echo cancellation, and wherein, using as reference signals a set of a plurality of  
5 low-correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals (e.g. suitably combining said multi-channel audio signals to  
10 produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals and using a set of said plurality of low-correlation composite signals as reference signals, or directly inputting a set of a plurality of low-  
15 correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals and using the set of said plurality of low-correlation composite  
20 signals as reference signals, or the like), composite transfer functions of said plurality of audio transfer systems are respectively derived, thereby to set corresponding filter characteristics. According to this invention, using as reference signals a set of a plurality of  
25 low-correlation composite signals which correspond to signals obtained by suitably combining multi-channel audio signals having a correlation therebetween and which have a lower

correlation with each other than that between such multi-channel audio signals, the composite transfer functions of said plurality of audio transfer systems are respectively derived, and corresponding filter characteristics are set, thereby to enable echo cancellation. In accordance therewith, since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the multi-channel audio signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation of respectively deriving the composite transfer functions of said plurality of audio transfer systems using as the reference signals the set of the plurality of low-correlation composite signals, may be, for example, a calculation of respectively deriving the composite transfer functions of said plurality of audio transfer systems based on a cross-spectrum calculation between the plurality of low-correlation composite signals and the composite signals of the individual collected audio signals of the respective microphones. Further, the calculation of respectively deriving the composite transfer functions of said plurality of audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of combining said multi-channel audio signals through addition or subtraction to produce a plurality of low-correlation composite signals

having a lower correlation with each other than that between said multi-channel audio signals, deriving cross spectra between said plurality of low-correlation composite signals and the composite signals of the individual collected audio signals of the respective microphones, and ensemble-averaging them in a predetermined time period per cross spectrum to derive composite transfer functions of said plurality of audio transfer systems.

[0038]

A multi-channel echo cancel method of this invention is a method wherein, with respect to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones and forming a plurality of audio transfer systems in which multi-channel sounds reproduced by said respective loudspeakers and having a correlation with each other are collected by said microphones, composite transfer functions of said plurality of audio transfer systems are estimated so as to set corresponding filter characteristics, respectively, echo cancel signals are respectively produced by giving said set filter characteristics to individual signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from composite signals of individual collected audio signals of said one or plurality of microphones, thereby performing echo cancellation, and wherein, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio

signals and which have a lower correlation with each other than that between said multi-channel audio signals (e.g. suitably combining said multi-channel audio signals to produce a plurality of low-correlation composite signals

5 having a lower correlation with each other than that between said multi-channel audio signals and using a set of said plurality of low-correlation composite signals as reference signals, or directly inputting a set of a plurality of low-correlation composite signals which correspond to signals

10 obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals and using the set of said plurality of low-correlation composite signals as reference signals, or the like), estimated errors

15 of composite transfer functions of said plurality of audio transfer systems are respectively derived, thereby to update corresponding filter characteristics to values that cancel said estimated errors. According to this invention, using as reference signals a set of a plurality of low-correlation

20 composite signals which correspond to signals obtained by suitably combining multi-channel audio signals having a correlation therebetween and which have a lower correlation with each other than that between such multi-channel audio signals, the estimated errors of the composite transfer

25 functions of said plurality of audio transfer systems are respectively derived so as to successively update the corresponding filter characteristics to the values that

cancel said estimated errors, thereby to enable echo  
cancellation. In accordance therewith, since the multi-  
channel audio signals can be reproduced from the loudspeakers  
with no or less processing, which induces deterioration,  
5 applied to the multi-channel audio signals, excellent  
reproduced tone quality can be achieved. Further, there is  
no or only a small delay in reproduced signals. Thus, when  
applying to the teleconferencing system or the like, natural  
conversation can be conducted. It is also possible to update  
10 the filter characteristics in real time. The calculation of  
respectively deriving the estimated errors of the composite  
transfer functions of said plurality of audio transfer  
systems using the set of the plurality of low-correlation  
composite signals as the reference signals, may be, for  
15 example, a calculation of respectively deriving the estimated  
errors of the composite transfer functions of said plurality  
of audio transfer systems based on a cross-spectrum  
calculation between said plurality of low-correlation  
composite signals and echo cancel error signals obtained by  
20 subtracting the corresponding echo cancel signals from the  
composite signals of the individual collected audio signals  
of said one or plurality of microphones. Further, the  
calculation of respectively deriving the estimated errors of  
the composite transfer functions of said plurality of audio  
25 transfer systems based on said cross-spectrum calculation,  
may be, for example, a calculation of combining said multi-  
channel audio signals through addition or subtraction to

produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals, deriving cross spectra between said plurality of low-correlation composite signals and the echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the composite signals of the individual collected audio signals of said one or plurality of microphones, and ensemble-averaging them in a predetermined time period per cross spectrum to derive estimated errors of the composite transfer functions of said plurality of audio transfer systems. Further, the correlation between said plurality of low-correlation composite signals is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0039]

A multi-channel sound transfer method of this invention is such that, with respect to two spaces each forming said plurality of audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the multi-channel audio signals, which have been echo-canceled by performing said method, are transmitted between said two spaces. In accordance therewith, the multi-channel audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be applied to the teleconferencing system or the like.

[0040]

A stereo echo cancel method of this invention is a method wherein, with respect to a space provided therein with two loudspeakers and one or two microphones and forming two  
5 or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said microphones, composite transfer functions of said two or four audio transfer systems are estimated so as to set corresponding filter characteristics, respectively, echo  
10 cancel signals are respectively produced by giving said set filter characteristics to individual signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from composite signals of individual collected audio signals of said one or two microphones, thereby  
15 performing echo cancellation, and wherein, using a sum signal and a difference signal of said stereo audio signals as reference signals, composite transfer functions of said two or four audio transfer systems are respectively derived, thereby to set corresponding filter characteristics.

20 According to this invention, since the sum signal and the difference signal of the stereo audio signals have a low correlation therebetween, estimated errors of the composite transfer functions of the two or four audio transfer systems are respectively derived using the sum signal and the  
25 difference signal as reference signals, so as to successively update the corresponding filter characteristics to values that cancel said estimated errors, thereby to enable echo

cancellation. In accordance therewith, since the stereo signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the stereo signals, excellent reproduced tone quality can be  
5 achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation of respectively deriving the composite transfer functions of said two or four audio  
10 transfer systems using the sum signal and the difference signal of said stereo audio signals as the reference signals, may be, for example, a calculation of respectively deriving the composite transfer functions of said two or four audio transfer systems based on a cross-spectrum calculation  
15 between the sum signal and the difference signal, and the composite signals of the individual collected audio signals of the respective microphones. Further, the calculation of respectively deriving the composite transfer functions of said two or four audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of  
20 deriving cross spectra between the sum signal and the difference signal of said stereo audio signals and the composite signals of the individual collected audio signals of the respective microphones, and ensemble-averaging them in  
25 a predetermined time period per cross spectrum to derive composite transfer functions of said two or four audio transfer systems.

[0041]

A stereo echo cancel method of this invention is a method wherein, with respect to a space provided therein with two loudspeakers and one or two microphones and forming two  
5 or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said microphones, composite transfer functions of said two or four audio transfer systems are estimated so as to set corresponding filter characteristics, respectively, echo  
10 cancel signals are respectively produced by giving said set filter characteristics to individual signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from composite signals of individual collected audio signals of said one or two microphones, thereby  
15 performing echo cancellation, and wherein, using a sum signal and a difference signal of said stereo audio signals as reference signals, estimated errors of composite transfer functions of said two or four audio transfer systems are respectively derived, thereby to update corresponding filter  
20 characteristics to values that cancel said estimated errors. According to this invention, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining multi-channel audio signals having a correlation therebetween and  
25 which have a lower correlation with each other than that between such multi-channel audio signals, the estimated errors of the composite transfer functions of said plurality

of audio transfer systems are respectively derived, so as to successively update the corresponding filter characteristics to the values that cancel the estimated errors, thereby to enable echo cancellation. In accordance therewith, since the  
5 multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the multi-channel audio signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals.

10 Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. It is also possible to update the filter characteristics in real time. The calculation of respectively deriving the estimated errors of the composite transfer functions of said two or four audio  
15 transfer systems using the sum signal and the difference signal of said stereo audio signals as the reference signals, may be, for example, a calculation of respectively deriving the estimated errors of the composite transfer functions of said two or four audio transfer systems based on a cross-  
20 spectrum calculation between the sum signal and the difference signal of said stereo audio signals and respective echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the composite signals of the individual collected audio signals of said one or two  
25 microphones. Further, the calculation of respectively deriving the estimated errors of the composite transfer functions of said two or four audio transfer systems based on

the cross-spectrum calculation between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, may be, for example, a calculation of deriving cross spectra between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, and ensemble-averaging them in a predetermined time period per cross spectrum to derive estimated errors of the composite transfer functions of said two or four audio transfer systems. Further, the correlation between the sum signal and the difference signal of said stereo audio signals is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0042]

A stereo audio transmission method of this invention is such that, with respect to two spaces each forming said four audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the stereo audio signals, which have been echo-canceled by performing said method, are transmitted between said two spaces. In accordance therewith, the stereo audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be applied to the teleconferencing system or the like.

[0043]

A stereo echo canceller of this invention is a

stereo echo canceller wherein, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are  
5 collected by said respective microphones, an audio signal supplied to the first loudspeaker is subjected to convolution calculations by first and second filter means, respectively, so as to produce first and second echo cancel signals, an audio signal supplied to the second loudspeaker is subjected  
10 to convolution calculations by third and fourth filter means, respectively, so as to produce third and fourth echo cancel signals, echo cancellation is performed by subtracting, using first subtracting means, said first and third echo cancel signals from a sum signal of collected audio signals of the  
15 respective microphones, and echo cancellation is performed by subtracting, using second subtracting means, said second and fourth echo cancel signals from a difference signal of the collected audio signals of the respective microphones, said stereo echo canceller comprising: transfer function  
20 calculating means for respectively deriving filter characteristics corresponding to composite transfer functions of said four audio transfer systems based on a cross-spectrum calculation between a sum signal and a difference signal of stereo audio signals to be reproduced by said respective  
25 loudspeakers and the sum signal and the difference signal of the respective microphone collected audio signals, thereby to set said derived filter characteristics to corresponding ones

of said first to fourth filter means, respectively.

[0044]

A stereo echo canceller of this invention is a stereo echo canceller wherein, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, an audio signal supplied to the first loudspeaker is subjected to convolution calculations by first and second filter means, respectively, so as to produce first and second echo cancel signals, an audio signal supplied to the second loudspeaker is subjected to convolution calculations by third and fourth filter means, respectively, so as to produce third and fourth echo cancel signals, echo cancellation is performed by subtracting, using first subtracting means, said first and third echo cancel signals from a sum signal of collected audio signals of the respective microphones, and echo cancellation is performed by subtracting, using second subtracting means, said second and fourth echo cancel signals from a difference signal of the collected audio signals of the respective microphones, said stereo echo canceller comprising: transfer function calculating means for respectively deriving estimated errors of composite transfer functions of said four audio transfer systems based on a cross-spectrum calculation between a sum signal and a difference signal of stereo audio signals to be reproduced by said respective loudspeakers and respective

echo cancel error signals obtained by subtracting the  
corresponding echo cancel signals from the sum signal and the  
difference signal of the respective microphone collected  
audio signals, thereby to update filter characteristics of  
5 said first to fourth filter means to values that cancel said  
estimated errors, respectively.

[0045]

The stereo echo canceller of this invention may be  
further provided with correlation detecting means for  
10 detecting the correlation between the sum signal and the  
difference signal of said stereo audio signals and, when a  
value of said correlation is no less than a prescribed value,  
stopping updating of said filter characteristics, thereby to  
prevent the echo cancel error signals from unexpectedly  
15 increasing.

[0046]

A stereo sound transfer apparatus of this invention  
is such that, with respect to two spaces each forming said  
four audio transfer systems, any of said stereo echo  
20 cancellers is arranged in each space, so that the stereo  
audio signals, which have been echo-canceled by said stereo  
echo cancellers, are transmitted between said two spaces.

[0047]

[Inventions of Claims 32 to 35 and Inventions  
25 relating to such Inventions]

A multi-channel echo cancel method of this invention  
is a method wherein, with respect to a space provided therein

with a plurality of loudspeakers and one or a plurality of  
microphones and forming a plurality of audio transfer systems  
in which multi-channel sounds reproduced by said respective  
loudspeakers and having a correlation with each other are  
5 collected by said microphones, composite transfer functions  
of said plurality of audio transfer systems are estimated so  
as to set corresponding filter characteristics, respectively,  
echo cancel signals are respectively produced by giving said  
set filter characteristics to composite signals of individual  
10 signals to be reproduced by said respective loudspeakers, and  
said echo cancel signals are subtracted from composite  
signals of individual collected audio signals of said one or  
plurality of microphones, thereby performing echo  
cancellation, and wherein, using as reference signals a set  
15 of a plurality of low-correlation composite signals which  
correspond to signals obtained by suitably combining said  
multi-channel audio signals and which have a lower  
correlation with each other than that between said multi-  
channel audio signals (e.g. suitably combining said multi-  
20 channel audio signals to produce a plurality of low-  
correlation composite signals having a lower correlation with  
each other than that between said multi-channel audio signals  
and using a set of said plurality of low-correlation  
composite signals as reference signals, or directly inputting  
25 a set of a plurality of low-correlation composite signals  
which correspond to signals obtained by suitably combining  
said multi-channel audio signals and which have a lower

correlation with each other than that between said multi-channel audio signals and using the set of said plurality of low-correlation composite signals as reference signals, or the like), composite transfer functions of said plurality of audio transfer systems are respectively derived, thereby to set corresponding filter characteristics. According to this invention, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining multi-channel audio signals having a correlation therebetween and which have a lower correlation with each other than that between such multi-channel audio signals, the composite transfer functions of said plurality of audio transfer systems are respectively derived, and corresponding filter characteristics are set, thereby to enable echo cancellation. In accordance therewith, since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the multi-channel audio signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation of respectively deriving the composite transfer functions of said plurality of audio transfer systems using as the reference signals the set of the plurality of low-correlation composite signals, may be, for example, a calculation of respectively deriving the composite transfer functions of

said plurality of audio transfer systems based on a cross-spectrum calculation between the plurality of low-correlation composite signals and the composite signals of the individual collected audio signals of the respective microphones.

5 Further, the calculation of respectively deriving the composite transfer functions of said plurality of audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of combining said multi-channel audio signals through addition or subtraction to  
10 produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals, deriving cross spectra between said plurality of low-correlation composite signals and the composite signals of the individual collected audio  
15 signals of the respective microphones, and ensemble-averaging them in a predetermined time period per cross spectrum to derive composite transfer functions of said plurality of audio transfer systems.

[0048]

20 A multi-channel echo cancel method of this invention is a method wherein, with respect to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones and forming a plurality of audio transfer systems in which multi-channel sounds reproduced by said respective  
25 loudspeakers and having a correlation with each other are collected by said microphones, composite transfer functions of said plurality of audio transfer systems are estimated so

as to set corresponding filter characteristics, respectively,  
echo cancel signals are respectively produced by giving said  
set filter characteristics to composite signals of individual  
signals to be reproduced by said respective loudspeakers, and  
5 said echo cancel signals are subtracted from composite  
signals of individual collected audio signals of said one or  
plurality of microphones, thereby performing echo  
cancellation, and wherein, using as reference signals a set  
of a plurality of low-correlation composite signals which  
10 correspond to signals obtained by suitably combining said  
multi-channel audio signals and which have a lower  
correlation with each other than that between said multi-  
channel audio signals (e.g. suitably combining said multi-  
channel audio signals to produce a plurality of low-  
15 correlation composite signals having a lower correlation with  
each other than that between said multi-channel audio signals  
and using a set of said plurality of low-correlation  
composite signals as reference signals, or directly inputting  
a set of a plurality of low-correlation composite signals  
20 which correspond to signals obtained by suitably combining  
said multi-channel audio signals and which have a lower  
correlation with each other than that between said multi-  
channel audio signals and using the set of said plurality of  
low-correlation composite signals as reference signals, or  
25 the like), estimated errors of composite transfer functions  
of said plurality of audio transfer systems are respectively  
derived, thereby to update corresponding filter

characteristics to values that cancel said estimated errors. According to this invention, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining multi-  
5 channel audio signals having a correlation therebetween and which have a lower correlation with each other than that between such multi-channel audio signals, the estimated errors of the composite transfer functions of said plurality of audio transfer systems are respectively derived so as to  
10 successively update the corresponding filter characteristics to the values that cancel said estimated errors, thereby to enable echo cancellation. In accordance therewith, since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces  
15 deterioration, applied to the multi-channel audio signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. It is also  
20 possible to update the filter characteristics in real time. The calculation of respectively deriving the estimated errors of the composite transfer functions of said plurality of audio transfer systems using the set of the plurality of low-correlation composite signals as the reference signals, may  
25 be, for example, a calculation of respectively deriving the estimated errors of the composite transfer functions of said plurality of audio transfer systems based on a cross-spectrum

calculation between said plurality of low-correlation composite signals and echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the composite signals of the individual collected audio signals of said one or plurality of microphones. Further, the calculation of respectively deriving the estimated errors of the composite transfer functions of said plurality of audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of combining said multi-channel audio signals through addition or subtraction to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals, deriving cross spectra between said plurality of low-correlation composite signals and the echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the composite signals of the individual collected audio signals of said one or plurality of microphones, and ensemble-averaging them in a predetermined time period per cross spectrum to derive estimated errors of the composite transfer functions of said plurality of audio transfer systems. Further, the correlation between said plurality of low-correlation composite signals is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0049]

A multi-channel sound transfer method of this invention is such that, with respect to two spaces each forming said plurality of audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the multi-channel audio signals, which have been echo-canceled by performing said method, are transmitted between said two spaces. In accordance therewith, the multi-channel audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be applied to the teleconferencing system or the like.

[0050]

A stereo echo cancel method of this invention is a method wherein, with respect to a space provided therein with two loudspeakers and one or two microphones and forming two or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said microphones, composite transfer functions of said two or four audio transfer systems are estimated so as to set corresponding filter characteristics, respectively; echo cancel signals are respectively produced by giving said set filter characteristics to composite signals of individual signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from composite signals of individual collected audio signals of said one or two microphones, thereby performing echo cancellation, and wherein, using a sum signal and a difference signal of said stereo audio signals as reference signals, composite transfer

functions of said two or four audio transfer systems are respectively derived, thereby to set corresponding filter characteristics. According to this invention, since the sum signal and the difference signal of the stereo audio signals have a low correlation therebetween, estimated errors of the composite transfer functions of the two or four audio transfer systems are respectively derived using the sum signal and the difference signal as reference signals, so as to successively update the corresponding filter characteristics to values that cancel said estimated errors, thereby to enable echo cancellation. In accordance therewith, since the stereo signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the stereo signals, excellent reproduced tone quality can be achieved. Further, there is no or only a small delay in reproduced signals: Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. The calculation of respectively deriving the composite transfer functions of said two or four audio transfer systems using the sum signal and the difference signal of said stereo audio signals as the reference signals, may be, for example, a calculation of respectively deriving the composite transfer functions of said two or four audio transfer systems based on a cross-spectrum calculation between the sum signal and the difference signal, and the composite signals of the individual collected audio signals of the respective

microphones. Further, the calculation of respectively deriving the composite transfer functions of said two or four audio transfer systems based on said cross-spectrum calculation, may be, for example, a calculation of deriving cross spectra between the sum signal and the difference  
5 signal of said stereo audio signals and the composite signals of the individual collected audio signals of the respective microphones, and ensemble-averaging them in a predetermined time period per cross spectrum to derive composite transfer  
10 functions of said two or four audio transfer systems.

[0051]

A stereo echo cancel method of this invention is a method wherein, with respect to a space provided therein with two loudspeakers and one or two microphones and forming two  
15 or four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said microphones, composite transfer functions of said two or four audio transfer systems are estimated so as to set corresponding filter characteristics, respectively, echo  
20 cancel signals are respectively produced by giving said set filter characteristics to composite signals of individual signals to be reproduced by said respective loudspeakers, and said echo cancel signals are subtracted from composite  
25 signals of individual collected audio signals of said one or two microphones, thereby performing echo cancellation, and wherein, using a sum signal and a difference signal of said stereo audio signals as reference signals, estimated errors

of composite transfer functions of said two or four audio transfer systems are respectively derived, thereby to update corresponding filter characteristics to values that cancel said estimated errors. According to this invention, using as  
5 reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining multi-channel audio signals having a correlation therebetween and which have a lower correlation with each other than that between such multi-channel audio  
10 signals, the estimated errors of the composite transfer functions of said plurality of audio transfer systems are respectively derived, so as to successively update the corresponding filter characteristics to the values that cancel the estimated errors, thereby to enable echo  
15 cancellation. In accordance therewith, since the multi-channel audio signals can be reproduced from the loudspeakers with no or less processing, which induces deterioration, applied to the multi-channel audio signals, excellent reproduced tone quality can be achieved. Further, there is  
20 no or only a small delay in reproduced signals. Thus, when applying to the teleconferencing system or the like, natural conversation can be conducted. It is also possible to update the filter characteristics in real time. The calculation of respectively deriving the estimated errors of the composite  
25 transfer functions of said two or four audio transfer systems using the sum signal and the difference signal of said stereo audio signals as the reference signals, may be, for example,

a calculation of respectively deriving the estimated errors of the composite transfer functions of said two or four audio transfer systems based on a cross-spectrum calculation between the sum signal and the difference signal of said stereo audio signals and respective echo cancel error signals obtained by subtracting the corresponding echo cancel signals from the composite signals of the individual collected audio signals of said one or two microphones. Further, the calculation of respectively deriving the estimated errors of the composite transfer functions of said two or four audio transfer systems based on the cross-spectrum calculation between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, may be, for example, a calculation of deriving cross spectra between the sum signal and the difference signal of said stereo audio signals and said echo cancel error signals, and ensemble-averaging them in a predetermined time period per cross spectrum to derive estimated errors of the composite transfer functions of said two or four audio transfer systems. Further, the correlation between the sum signal and the difference signal of said stereo audio signals is detected and, when a value of said correlation is no less than a prescribed value, updating of said filter characteristics is stopped, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0052]

A stereo audio transmission method of this invention

is such that, with respect to two spaces each forming said four audio transfer systems, any of the foregoing multi-channel echo cancel methods is carried out, so that the stereo audio signals, which have been echo-canceled by performing said method, are transmitted between said two spaces. In accordance therewith, the stereo audio transmission with reduced echo cancellation can be performed between two spots, which, for example, can be applied to the teleconferencing system or the like.

[0053]

A stereo echo canceller of this invention is a stereo echo canceller wherein, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, a sum signal of stereo audio signals to be reproduced by said respective loudspeakers is subjected to convolution calculations by first and second filter means, respectively, so as to produce first and second echo cancel signals, a difference signal of the stereo audio signals to be reproduced by said respective loudspeakers is subjected to convolution calculations by third and fourth filter means, respectively, so as to produce third and fourth echo cancel signals, echo cancellation is performed by subtracting, using first subtracting means, said first and third echo cancel signals from a sum signal of collected audio signals of the respective microphones, and

echo cancellation is performed by subtracting, using second subtracting means, said second and fourth echo cancel signals from a difference signal of the collected audio signals of the respective microphones, said stereo echo canceller comprising: transfer function calculating means for respectively deriving filter characteristics corresponding to composite transfer functions of said four audio transfer systems based on a cross-spectrum calculation between the sum signal and the difference signal of the stereo audio signals to be reproduced by said respective loudspeakers and the sum signal and the difference signal of the respective microphone collected audio signals, thereby to set said derived filter characteristics to corresponding ones of said first to fourth filter means, respectively.

[0054]

A stereo echo canceller of this invention is a stereo echo canceller wherein, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, a sum signal of stereo audio signals to be reproduced by said respective loudspeakers is subjected to convolution calculations by first and second filter means, respectively, so as to produce first and second echo cancel signals, a difference signal of the stereo audio signals to be reproduced by said respective loudspeakers is subjected to convolution calculations by

third and fourth filter means, respectively, so as to produce  
third and fourth echo cancel signals, echo cancellation is  
performed by subtracting, using first subtracting means, said  
first and third echo cancel signals from a sum signal of  
5 collected audio signals of the respective microphones, and  
echo cancellation is performed by subtracting, using second  
subtracting means, said second and fourth echo cancel signals  
from a difference signal of the collected audio signals of  
the respective microphones, said stereo echo canceller  
10 comprising: transfer function calculating means for  
respectively deriving estimated errors of composite transfer  
functions of said four audio transfer systems based on a  
cross-spectrum calculation between the sum signal and the  
difference signal of the stereo audio signals to be  
15 reproduced by said respective loudspeakers and respective  
echo cancel error signals obtained by subtracting the  
corresponding echo cancel signals from the sum signal and the  
difference signal of the respective microphone collected  
audio signals, thereby to update filter characteristics of  
20 said first to fourth filter means to values that cancel said  
estimated errors, respectively.

[0055]

The stereo echo canceller of this invention may be  
further provided with correlation detecting means for  
25 detecting the correlation between the sum signal and the  
difference signal of said stereo audio signals and, when a  
value of said correlation is no less than a prescribed value,

stopping updating of said filter characteristics, thereby to prevent the echo cancel error signals from unexpectedly increasing.

[0056]

5           A stereo sound transfer apparatus of this invention is such that, with respect to two spaces each forming said four audio transfer systems, any of said stereo echo cancellers is arranged in each space, so that the stereo audio signals, which have been echo-canceled by said stereo  
10   echo cancellers, are transmitted between said two spaces.

[0057]

[Inventions of Claims 36 to 40 and Inventions  
relating to such Inventions]

15           A transfer function calculation apparatus of this invention is a transfer function calculation apparatus which, with respect to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones and forming a plurality of audio transfer systems in which multi-channel sounds inputted from an outside and reproduced by  
20   said respective loudspeakers and having a correlation with each other are collected by said microphones, estimates individual transfer functions of said plurality of audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual  
25   transfer functions, wherein, using as reference signals a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining said

multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals (e.g. suitably combining said multi-channel audio signals to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals and using a set of said plurality of low-correlation composite signals as reference signals, or directly inputting a set of a plurality of low-correlation composite signals which correspond to signals obtained by suitably combining said multi-channel audio signals and which have a lower correlation with each other than that between said multi-channel audio signals and using the set of said plurality of low-correlation composite signals as reference signals, or the like), individual transfer functions of the respective audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions are estimated. The calculation of respectively deriving the individual transfer functions of the respective audio transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual transfer functions, using as the reference signals the set of the plurality of low-correlation composite signals, may be, for example, a calculation of respectively deriving the individual transfer functions of the respective audio transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual

transfer functions, based on a cross-spectrum calculation between the plurality of low-correlation composite signals and the individual collected audio signals of the microphones, or the plurality of composite signals obtained by suitably combining said individual collected audio signals. Further, the calculation of respectively deriving the individual transfer functions of said plurality of audio transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual transfer functions, based on said cross-spectrum calculation, may be, for example, a calculation of respectively deriving the individual transfer functions of said plurality of audio transfer systems or the plurality of composite transfer functions obtained by suitably combining said individual transfer functions, by combining said multi-channel audio signals through addition or subtraction to produce a plurality of low-correlation composite signals having a lower correlation with each other than that between said multi-channel audio signals, deriving cross spectra between said plurality of low-correlation composite signals and the individual collected audio signals of the microphones, or the plurality of composite signals obtained by suitably combining said individual collected audio signals, and ensemble-averaging them in a predetermined time period per cross spectrum. The calculation of respectively deriving the individual transfer functions of said plurality of audio transfer systems based on said cross-spectrum calculation may

also be a calculation of respectively deriving the individual transfer functions of said plurality of audio transfer systems by producing a plurality of mutually orthogonal uncorrelated composite signals by applying a principal component analysis to said multi-channel audio signals, deriving cross spectra between said plurality of uncorrelated composite signals and the individual collected audio signals of the microphones, and ensemble-averaging them in a predetermined time period per cross spectrum.

10 [0058]

A transfer function calculation apparatus of this invention is a transfer function calculation apparatus which, with respect to a space provided therein with two loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, estimates individual transfer functions of said four audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions, wherein, using a sum signal and a difference signal of said stereo audio signals as reference signals, individual transfer functions of said four audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions are estimated. The calculation of respectively deriving the individual transfer functions of said four audio transfer systems or the plurality of

composite transfer functions obtained by suitably combining  
said individual transfer functions, using the sum signal and  
the difference signal of said stereo audio signals as the  
reference signals, may be, for example, a calculation of  
5 respectively deriving the individual transfer functions of  
said four audio transfer systems or the plurality of  
composite transfer functions obtained by suitably combining  
said individual transfer functions, based on a cross-spectrum  
calculation between said sum signal and said difference  
10 signal, and individual collected audio signals of the  
microphones, or a plurality of composite signals obtained by  
suitably combining said individual collected audio signals.  
Further, the calculation of respectively deriving the  
individual transfer functions of said four audio transfer  
15 systems or the plurality of composite transfer functions  
obtained by suitably combining said individual transfer  
functions, based on said cross-spectrum calculation, may be a  
calculation of deriving cross spectra between the sum signal  
and the difference signal of said stereo audio signals and  
20 the individual collected audio signals of the microphones, or  
the plurality of composite signals obtained by suitably  
combining said individual collected audio signals, and  
ensemble-averaging them in a predetermined time period per  
cross spectrum, thereby to respectively derive the individual  
25 transfer functions of said four audio transfer systems or the  
plurality of composite transfer functions obtained by  
suitably combining said individual transfer functions.

[0059]

A transfer function calculation apparatus of this invention is a transfer function calculation apparatus which, with respect to a space provided therein with two  
5 loudspeakers and two microphones and forming four audio transfer systems in which stereo sounds reproduced by said respective loudspeakers are collected by said respective microphones, estimates individual transfer functions of said four audio transfer systems, wherein mutually orthogonal two  
10 uncorrelated composite signals are produced by applying a principal component analysis to said stereo audio signals, and individual transfer functions of said four audio transfer systems are estimated using a set of said two uncorrelated composite signals as reference signals. The calculation of  
15 respectively deriving the individual transfer functions of said four audio transfer systems using said two uncorrelated composite signals as the reference signals, may be, for example, a calculation of respectively deriving the individual transfer functions of said four audio transfer  
20 systems based on a cross-spectrum calculation between said two uncorrelated composite signals and the individual collected audio signals of the microphones. Further, the calculation of respectively deriving the individual transfer functions of said four audio transfer systems based on said  
25 cross-spectrum calculation, may be, for example, a calculation of deriving cross spectra between said two uncorrelated composite signals and the individual collected

audio signals of the microphones, and ensemble-averaging them in a predetermined time period per cross spectrum, thereby to respectively derive the individual transfer functions of said four audio transfer systems. In this case, by making

5 relatively longer a time period of performing said ensemble averaging when double talk is detected where sounds other than those reproduced by said loudspeakers are inputted into said microphones, while making it relatively shorter when the double talk is not detected, it is possible to fully converge  
10 the estimated errors when the double talk exists, and further, quicken the convergence of the estimated errors when there is no double talk.

[0060]

[Inventions of Claims 41 to 51 and Inventions  
15 relating to such Inventions]

An inventive echo cancel method is associated to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones for forming a plurality of audio transfer systems through which audio signals of multi-  
20 channels having a correlation with each other are reproduced by said respective loudspeakers and are collected by said microphones, and designed for performing an echo cancellation by subtracting an echo cancel signal from the audio signals collected by the respective microphone or from composite  
25 signals obtained by combining the collected audio signals. The inventive method comprises inputting a plurality of low-correlation audio signals which are obtained by suitably

combining first audio signals of multi-channels and which have a lower correlation with each other than that among said first audio signals of multi-channels, generating second audio signals of multi-channels having a correlation with each other by computation based on the inputted low-correlation audio signals, feeding the generated second audio signals to the respective loudspeakers so as to reproduce audio sounds, feeding the generated second audio signals or the inputted low-correlation audio signals to filters, estimating individual transfer functions of said plurality of said audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions based on the inputted low-correlation audio signals so as to set corresponding filter characteristics, producing echo cancel signals by applying said set filter characteristics to the second audio signals or the low-correlation audio signals fed to the filters, and subtracting said echo cancel signals from collected audio signals obtained by collecting the reproduced audio sounds by the microphones or from composite audio signals obtained by suitably combining said collected audio signals, thereby performing the echo cancellation.

Preferably, in the inventive echo cancel method, the inputted low-correlation audio signals are obtained by adding or subtracting the first audio signals of multi-channels with each other.

Another inventive echo cancel method is associated

to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones for forming a plurality of audio transfer systems through which audio signals of multi-channels having a correlation with each other are reproduced by said respective loudspeakers and are collected by said microphones, and designed for performing an echo cancellation by subtracting an echo cancel signal from the audio signals collected by the respective microphone or from composite signals obtained by combining the collected audio signals. The inventive method comprises inputting a plurality of first low-correlation audio signals which are obtained by suitably combining first audio signals of multi-channels and which have a lower correlation with each other than that among said first audio signals of multi-channels, generating second audio signals of multi-channels having a correlation with each other by computation based on the inputted first low-correlation audio signals, feeding the generated second audio signals to the respective loudspeakers so as to reproduce audio sounds, generating second low-correlation audio signals of multi-channels based on the generated second audio signals, feeding the generated second audio signals or the generated second low-correlation audio signals to filters, estimating individual transfer functions of said plurality of said audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions based on the generated second low-correlation audio signals so as

to set corresponding filter characteristics, producing echo cancel signals by applying said set filter characteristics to the second audio signals or the second low-correlation audio signals fed to the filters, and subtracting said echo cancel  
5 signals from collected audio signals obtained by collecting the reproduced audio sounds at the microphones or from composite audio signals obtained by suitably combining said collected audio signals, thereby performing the echo cancellation.

10            Preferably, in the inventive echo cancel method, the inputted first low-correlation audio signals are obtained by adding or subtracting the first audio signals of multi-channels with each other.

15            An inventive echo canceller is associated to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones for forming a plurality of audio transfer systems through which audio signals of multi-channels having a correlation with each other are reproduced by said respective loudspeakers and are collected by said  
20 microphones, and designed for performing an echo cancellation by subtracting an echo cancel signal from the audio signals collected by the respective microphone or from composite signals obtained by combining the collected audio signals. The inventive echo canceller comprises an inputting means for  
25 inputting a plurality of low-correlation audio signals which are obtained by suitably combining first audio signals of multi-channels and which have a lower correlation with each

other than that among said first audio signals of multi-channels, a demodulating means provided for generating second audio signals of multi-channels having a correlation with each other by demodulating the inputted low-correlation audio signals, and for feeding the generated second audio signals to the respective loudspeakers so as to reproduce audio sounds, an estimating means for estimating individual transfer functions of said plurality of said audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions based on the inputted low-correlation audio signals so as to set corresponding filter characteristics, a filter means for producing echo cancel signals by applying said set filter characteristics to the second audio signals or the low-correlation audio signals fed to the filter means, and a subtracting means for subtracting said echo cancel signals from collected audio signals obtained by collecting the reproduced audio sounds at the microphones or from composite audio signals obtained by suitably combining said collected audio signals, thereby performing the echo cancellation.

Preferably, in the inventive echo canceller, the inputted low-correlation audio signals are obtained by adding or subtracting the first audio signals of multi-channels with each other.

Another inventive echo canceller is associated to a space provided therein with a plurality of loudspeakers and one or a plurality of microphones for forming a plurality of

audio transfer systems through which audio signals of multi-channels having a correlation with each other are reproduced by said respective loudspeakers and are collected by said microphones, and designed for performing an echo cancellation  
5 by subtracting an echo cancel signal from the audio signals collected by the respective microphone or from composite signals obtained by combining the collected audio signals. The inventive echo canceller comprises an inputting means for inputting a plurality of first low-correlation audio signals  
10 which are obtained by suitably combining first audio signals of multi-channels and which have a lower correlation with each other than that among said first audio signals of multi-channels, a demodulating means provided for generating second audio signals of multi-channels having a correlation with  
15 each other by demodulating the inputted first low-correlation audio signals, and for feeding the generated second audio signals to the respective loudspeakers so as to reproduce audio sounds, an estimating means provided for generating second low-correlation audio signals of multi-channels based  
20 on the generated second audio signals, and for estimating individual transfer functions of said plurality of said audio transfer systems or a plurality of composite transfer functions obtained by suitably combining said individual transfer functions based on the generated second low-  
25 correlation audio signals so as to set corresponding filter characteristics, a filter means for producing echo cancel signals by applying said set filter characteristics to the

second audio signals or the second low-correlation audio signals fed to the filter means, and a subtracting means for subtracting said echo cancel signals from collected audio signals obtained by collecting the reproduced audio sounds at  
5 the microphones or from composite audio signals obtained by suitably combining said collected audio signals, thereby performing the echo cancellation.

Preferably, in the inventive echo canceller, the inputted first low-correlation audio signals are obtained by  
10 adding or subtracting the first audio signals of multi-channels with each other.

Preferably, in a inventive multi-channel echo canceller, the multi-channel audio signals being inputted from an outside and having a correlation with each other are  
15 reproduced by said respective loudspeakers without lowering the correlation of the inputted multi-channel audio signals.

Preferably, the multi-channel audio signals being inputted from an outside and having a correlation with each other are provisionally modulated to lower the correlation,  
20 then demodulated to restore the correlation, and thereafter reproduced by said respective loudspeakers. Further, the multi-channel audio signals are provisionally modulated to lower the correlation by adding and subtracting the multi-channel audio signals with each other, or by orthogonalizing  
25 the multi-channel audio signals with each other.

In this invention, "audio signal" is not limited to human voices, but covers all acoustic signals in the audible

frequency band.

#### Brief Description of the Drawings

Fig. 1 is a block diagram showing a structural  
5 example in a stereo echo canceller 16, 24 of Fig. 2.

Fig. 2 is a block diagram showing an embodiment of a  
stereo sound transfer apparatus of this invention.

Fig. 3 is a diagram showing the simulation  
measurement results of the echo cancel performance of the  
10 stereo echo canceller 16, 24 of Fig. 1.

Fig. 4 is a diagram showing the simulation  
measurement results of the echo cancel performance of the  
stereo echo canceller 16, 24 of Fig. 1.

Fig. 5 is a diagram showing the simulation  
15 measurement results of the echo cancel performance of the  
stereo echo canceller 16, 24 of Fig. 1.

Fig. 6 is a diagram showing the simulation  
measurement results of the echo cancel performance of the  
stereo echo canceller 16, 24 of Fig. 1.

20 Fig. 7 is a diagram showing the simulation  
measurement results of the echo cancel performance of the  
stereo echo canceller 16, 24 of Fig. 1.

Fig. 8 is a diagram showing the simulation  
measurement results of the echo cancel performance of the  
25 stereo echo canceller 16, 24 of Fig. 1.

Fig. 9 is a block diagram showing another structural  
example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 10 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 11 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 12 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 13 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 14 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 15 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 16 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 17 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 18 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 19 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 20 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 21 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 22 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 23 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 24 is a block diagram showing a modification of the structures of Figs. 1, 9 to 11.

Fig. 25 is a block diagram showing a modification of the structures of Figs. 12 to 15.

Fig. 26 is a block diagram showing a modification of the structures of Figs. 16 to 19.

Fig. 27 is a block diagram showing a modification of the structures of Figs. 20 and 21.

Fig. 28 is a block diagram showing another structural example in the stereo echo canceller 16, 24 of Fig. 2.

Fig. 29 is a time chart showing an example of unit

intervals for orthogonalization processing and deriving an impulse response or its estimated error in the stereo echo canceller of Fig. 28.

Fig. 30 is a diagram showing one example of  
5 functional blocks of an orthogonalizing filter 500 of Fig. 28.

Fig. 31 is a diagram showing one example of functional blocks of transfer function calculating means 502 of Fig. 28.

Fig. 32 is a functional block diagram showing a  
10 modification of the transfer function calculating means 502 of Fig. 31.

Fig. 33 is a diagram showing the simulation measurement result with respect to a time-domain variation in echo cancellation amount of the stereo echo canceller 16, 24  
15 of Fig. 28 when there is no double talk.

Fig. 34 is a diagram showing the simulation measurement result with respect to a time-domain variation in echo cancellation amount of the stereo echo canceller 16, 24 of Fig. 28 when there is no double talk.

20 Fig. 35 is a diagram showing the simulation measurement result with respect to a time-domain variation in echo cancellation amount of the stereo echo canceller 16, 24 of Fig. 28 when there is no double talk.

Fig. 36 is a diagram showing the simulation measurement result with respect to a time-domain variation in  
25 echo cancellation amount of the stereo echo canceller 16, 24 of Fig. 28 when there is no double talk.

Fig. 37 is a diagram showing the simulation measurement result with respect to a time-domain variation in echo cancellation amount of the stereo echo canceller 16, 24 of Fig. 28 when there is double talk.

5            Fig. 38 is a diagram showing the simulation measurement result with respect to a time-domain variation in echo cancellation amount of the stereo echo canceller 16, 24 of Fig. 28 when there is double talk.

10           Fig. 39 is a diagram showing the simulation measurement result with respect to a time-domain variation in echo cancellation amount of the stereo echo canceller 16, 24 of Fig. 28 when there is double talk.

15           Fig. 40 is a diagram showing the simulation measurement result with respect to a time-domain variation in echo cancellation amount of the stereo echo canceller 16, 24 of Fig. 28 when there is double talk.

20           Fig. 41 is a diagram showing the simulation measurement result with respect to a time-domain variation in transfer function estimated error of the stereo echo canceller 16, 24 of Fig. 28 when there is no double talk.

Fig. 42 is a diagram showing the simulation measurement result with respect to a time-domain variation in transfer function estimated error of the stereo echo canceller 16, 24 of Fig. 28 when there is no double talk.

25           Fig. 43 is a diagram showing the simulation measurement result with respect to a time-domain variation in transfer function estimated error of the stereo echo

canceller 16, 24 of Fig. 28 when there is no double talk.

Fig. 44 is a diagram showing the simulation measurement result with respect to a time-domain variation in transfer function estimated error of the stereo echo

5 canceller 16, 24 of Fig. 28 when there is no double talk.

Fig. 45 is a diagram showing the simulation measurement result with respect to a time-domain variation in transfer function estimated error of the stereo echo canceller 16, 24 of Fig. 28 when there is double talk.

10 Fig. 46 is a diagram showing the simulation measurement result with respect to a time-domain variation in transfer function estimated error of the stereo echo canceller 16, 24 of Fig. 28 when there is double talk.

15 Fig. 47 is a diagram showing the simulation measurement result with respect to a time-domain variation in transfer function estimated error of the stereo echo canceller 16, 24 of Fig. 28 when there is double talk.

20 Fig. 48 is a diagram showing the simulation measurement result with respect to a time-domain variation in transfer function estimated error of the stereo echo canceller 16, 24 of Fig. 28 when there is double talk.

Fig. 49 is an exemplary diagram for explaining an ensemble averaging process according to overlap processing.

25 Fig. 50 is a block diagram showing a modification of the stereo echo canceller 16, 24 of Fig. 28.

Fig. 51 is a block diagram showing a structural example wherein the number of microphones is modified to one

in Fig. 1.

## Best Mode for Carrying Out the Invention

[0061]

5           Embodiments of the present invention will be described hereinbelow. Fig. 2 shows the whole structure of a two-way stereo sound transfer apparatus according to this invention. This is for performing two-way stereo transmission between a spot A and a spot B and, for example,  
10           is applicable to the teleconferencing system. In the spot A, two loudspeakers SP-A(L) and SP-A(R) and two microphones MC-A(L) and MC-A(R) are arranged in one space. Collected audio signals of the microphones MC-A(L) and MC-A(R) are converted into digital signals at A/D converters 12 and 14,  
15           respectively, and applied with echo cancel processing at a stereo echo canceller 16, then modulated at a CODEC (CODER and DECODER) 18 and transmitted to the spot B via a wire or radio transmission line 20. In the spot B, two loudspeakers SP-B(L) and SP-B(R) and two microphones MC-B(L) and MC-B(R)  
20           are arranged in one space. Incidentally, assuming a situation that participants in the spots A and B talk with each other in a face-to-face manner in the teleconferencing system, loudspeakers and microphones are arranged such that a sound collected by a microphone on the participant's left in  
25           one spot is reproduced from a loudspeaker on the participant's right in the other spot whereas a sound collected by a microphone on the participant's right in one

spot is reproduced from a loudspeaker on the participant's left in the other spot. Specifically, when the loudspeaker SP-A(L) and the microphone MC-A(L) are arranged on the participant's left and the loudspeaker SP-A(R) and the microphone MC-A(R) are arranged on the participant's right in the spot A, the loudspeaker SP-B(L) and the microphone MC-B(L) are arranged on the participant's right and the loudspeaker SP-B(R) and the microphone MC-B(R) are arranged on the participant's left in the spot B.

10 [0062]

The signals transmitted from the spot A are inputted into a CODEC 22 where the collected audio signals of the microphones MC-A(L) and MC-A(R) are demodulated. These demodulated collected audio signals of the microphones MC-A(L) and MC-A(R) are respectively converted into analog signals at D/A converters 26 and 28 via a stereo echo canceller 24, and respectively reproduced at the loudspeakers SP-B(L) and SP-B(R). Collected audio signals of the microphones MC-B(L) and MC-B(R) in the spot B are converted into digital signals at A/D converters 30 and 32, respectively, and applied with echo cancel processing at the stereo echo canceller 24, then modulated at the CODEC 22 and transmitted to the spot A via the transmission line 20. The signals transmitted to the spot A are inputted into the CODEC 18 where the collected audio signals of the microphones MC-B(L) and MC-B(R) are demodulated. These demodulated collected audio signals of the microphones MC-B(L) and MC-

B(R) are respectively converted into analog signals at D/A converters 34 and 36 via the stereo echo canceller 16, and respectively reproduced at the loudspeakers SP-A(L) and SP-A(R).

5 [0063]

Fig. 1 shows a structural example in the stereo echo canceller 16, 24. Left/right two-channel stereo signals  $x_L$  and  $x_R$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are outputted  
10 from sound output ends SO(L) and SO(R) as they are (i.e. not through sum/difference signal producing means 52), and reproduced at loudspeakers SP(L) {representing SP-A(L) or SP-B(L)} and SP(R) {representing SP-A(R) or SP-B(R)}, respectively.

15 [0064]

Filer means 40-1 to 40-4 are formed by, for example, FIR filters. Of them, the filter means 40-1 is set with an impulse response corresponding to a transfer function between the loudspeaker SP(L) and a microphone MC(L) {representing  
20 MC-A(L) or MC-B(L)} and performs, using such an impulse response, a convolution calculation of a signal to be outputted from the sound output end SO(L), thereby producing an echo cancel signal EC1 corresponding to a signal obtained such that the signal outputted from the sound output end  
25 SO(L) is reproduced at the loudspeaker SP(L), collected by the microphone MC(L) and inputted into a sound input end SI(L). The filter means 40-2 is set with an impulse response

corresponding to a transfer function between the loudspeaker SP(L) and a microphone MC(R) {representing MC-A(R) or MC-B(R)} and performs, using such an impulse response, a convolution calculation of a signal to be outputted from the sound output end SO(L), thereby producing an echo cancel signal EC2 corresponding to a signal obtained such that the signal outputted from the sound output end SO(L) is reproduced at the loudspeaker SP(L), collected by the microphone MC(R) and inputted into a sound input end SI(R).

The filter means 40-3 is set with an impulse response corresponding to a transfer function between the loudspeaker SP(R) and the microphone MC(L) and performs, using such an impulse response, a convolution calculation of a signal to be outputted from the sound output end SO(R), thereby producing an echo cancel signal EC3 corresponding to a signal obtained such that the signal outputted from the sound output end SO(R) is reproduced at the loudspeaker SP(R), collected by the microphone MC(L) and inputted into the sound input end SI(L). The filter means 40-4 is set with an impulse response corresponding to a transfer function between the loudspeaker SP(R) and the microphone MC(R) and performs, using such an impulse response, a convolution calculation of a signal to be outputted from the sound output end SO(R), thereby producing an echo cancel signal EC4 corresponding to a signal obtained such that the signal outputted from the sound output end SO(R) is reproduced at the loudspeaker SP(R), collected by the microphone MC(R) and inputted into the sound input end

SI(R).

[0065]

An adder 44 performs a calculation of  $EC1+EC3$ . An adder 46 performs a calculation of  $EC2+EC4$ . A subtracter 48  
5 subtracts an echo cancel signal  $EC1+EC3$  from a collected audio signal of the microphone MC(L) inputted from the sound input end SI(L), thereby to perform echo cancellation. A subtracter 50 subtracts an echo cancel signal  $EC2+EC4$  from a collected audio signal of the microphone MC(R) inputted from  
10 the sound input end SI(R), thereby to perform echo cancellation. These echo-canceled signals of the respective left and right channels are outputted from line output ends LO(L) and LO(R), respectively, and transmitted toward the spot on the counterpart side.

15 [0066]

The sum/difference signal producing means 52 performs addition, using an adder 54, of the left/right two-channel stereo signals  $x_L$  and  $x_R$  inputted into the line input ends LI(L) and LI(R) so as to produce a sum signal  $x_L+x_R$ ,  
20 while performs subtraction thereof using a subtracter 56 so as to produce a difference signal  $x_L-x_R$  (or it may also be  $x_R-x_L$ ). In case of the left/right two-channel stereo signals, the sum signal  $x_L+x_R$  and the difference signal  $x_L-x_R$  are in general low in correlation therebetween, and frequently,  
25 approximately uncorrelated. Transfer function calculating means 58 implements a cross-spectrum calculation between the sum signal  $x_L+x_R$  and the difference signal  $x_L-x_R$  produced by

the sum/difference signal producing means 52 and signals  $e_L$  and  $e_R$  outputted from the subtracters 48 and 50 and, based on this cross-spectrum calculation, sets filter characteristics (impulse responses) of the filter means 40-1 to 40-4.

5 Specifically, upon starting the system, the filter characteristics of the filter means 40-1 to 40-4 are not set, i.e. coefficients are all set to zero, so that the echo cancel signals EC1 to EC4 are zero, and thus the collected audio signals of the microphones MC(L) and MC(R) themselves  
10 are outputted from the subtracters 48 and 50. Therefore, at this time, the transfer function calculating means 58 performs the cross-spectrum calculation between the sum signal  $x_L + x_R$  and the difference signal  $x_L - x_R$  produced by the sum/difference signal producing means 52 and the collected  
15 audio signals  $e_L$  and  $e_R$  of the microphones MC(L) and MC(R) outputted from the subtracters 48 and 50 and, based on this cross-spectrum calculation, derives transfer functions of four audio transfer systems between the loudspeakers SP(L) and SP(R) and the microphones MC(L) and MC(R), respectively,  
20 and implements initial setting of the filter characteristics of the filter means 40-1 to 40-4 to values corresponding to such transfer functions. After the initial setting, since the echo cancel signals are produced by the filter means 40-1 to 40-4, the echo cancel error signals  $e_L$  and  $e_R$   
25 corresponding to difference signals between the collected audio signals of the microphones MC(L) and MC(R) and the echo cancel signals EC1 to EC4 are outputted from the subtracters

48 and 50. Therefore, at this time, the transfer function calculating means 58 performs the cross-spectrum calculation between the sum signal  $x_L+x_R$  and the difference signal  $x_L-x_R$  produced by the sum/difference signal producing means 52 and the echo cancel error signals  $e_L$  and  $e_R$  outputted from the subtracters 48 and 50 and, based on this cross-spectrum calculation, derives estimated errors of the transfer functions of the four audio transfer systems between the loudspeakers SP(L) and SP(R) and the microphones MC(L) and MC(R), respectively, and updates the filter characteristics of the filter means 40-1 to 40-4 to values that cancel the estimated errors, respectively. By repeating this updating operation per prescribed time period, the echo cancel error can be converged to a minimum value. Further, even if the transfer functions change due to movement of the microphone positions or the like, the echo cancel error can be converged to a minimum value by sequentially updating the filter characteristics of the filter means 40-1 to 40-4 depending thereon.

[0067]

Correlation detecting means 60 detects a correlation between the sum signal  $x_L+x_R$  and the difference signal  $x_L-x_R$  based on a correlation value calculation or the like, and stops updating of the foregoing filter characteristics when the correlation value is no less than a prescribed value. When the correlation value becomes lower than the prescribed value, updating of the foregoing filter characteristics is

restarted. Incidentally, as a concrete technique for deriving the correlation between the sum signal  $x_L+x_R$  and the difference signal  $x_L-x_R$ , any of the known techniques for detecting a correlation of two signals may be used.

5 [0068]

Herein, the filter characteristics (impulse responses) that are set to the filter means 40-1 to 40-4 by the transfer function calculating means 58 will be described. In the following description, the transfer functions and the  
10 filter characteristics are expressed using the following symbols.

- $H_{xx}$  : a transfer function (frequency-axis expression)
- $h_{xx}$  : an impulse response (time-axis expression)  
15 corresponding to  $H_{xx}$
- $\hat{H}_{xx}$  : an estimated transfer function (a transfer function set to a filter)
- $\hat{h}_{xx}$  : an impulse response corresponding to  $\hat{H}_{xx}$
- $\Delta H_{xx}$  : a transfer function estimated error
- $\Delta h_{xx}$  : an impulse response corresponding to  $\Delta H_{xx}$   
20 (note: Suitable symbols are allocated to xx.)

The sum signal and the difference signal are respectively defined as follows.

sum signal :  $x_M (= x_L+x_R)$   
25 difference signal :  $x_S (= x_L-x_R)$

The transfer functions of the four audio transfer systems between the loudspeakers SP(L) and SP(R) and the

microphones MC(L) and MC(R) are respectively defined as follows.

$H_{LL}$ : a transfer function of the system from the loudspeaker SP(L) to the microphone MC(L)

5  $H_{LR}$ : a transfer function of the system from the loudspeaker SP(L) to the microphone MC(R)

$H_{RL}$ : a transfer function of the system from the loudspeaker SP(R) to the microphone MC(L)

10  $H_{RR}$ : a transfer function of the system from the loudspeaker SP(R) to the microphone MC(R)

The input signals  $x_L$  and  $x_R$  at the line input ends LI(L) and LI(R) of the stereo echo canceller 16, 24 are replaced by

$$x_L = (x_M + x_S) / 2$$

15  $x_R = (x_M - x_S) / 2.$

[0069]

Then, a calculation shown below is carried out in the transfer function calculating means 58. The signals  $e_L$  and  $e_R$  {the collected audio signals of the microphones MC(L) and MC(R) as they are before the initial setting of the filter means 40-1 to 40-4 whereas the echo cancel error signals after the initial setting} outputted from the subtracters 48 and 50, assuming that frequency-axis expressions of the signals  $x_M$ ,  $x_S$ ,  $e_L$  and  $e_R$  are respectively  
20 given as  $X_M$ ,  $X_S$ ,  $E_L$  and  $E_R$ , become

$$E_L = \{ (X_M + X_S) H_{LL} / 2 \} + \{ (X_M - X_S) H_{RL} / 2 \}$$

$$E_R = \{ (X_M + X_S) H_{LR} / 2 \} + \{ (X_M - X_S) H_{RR} / 2 \}$$

hence

$$\begin{aligned} 2E_L &= (X_M + X_S) H_{LL} + (X_M - X_S) H_{RL} \\ &= X_M (H_{LL} + H_{RL}) + X_S (H_{LL} - H_{RL}) \dots (1) \end{aligned}$$

5

$$\begin{aligned} 2E_R &= (X_M + X_S) H_{LR} + (X_M - X_S) H_{RR} \\ &= X_M (H_{LR} + H_{RR}) + X_S (H_{LR} - H_{RR}) \dots (2) \end{aligned}$$

When both sides of the equation (1) are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$  (i.e. deriving cross spectra) and ensemble-averaged,

$$\Sigma X_M^* \cdot 2E_L = \Sigma X_M^* \cdot X_M (H_{LL} + H_{RL}) + \Sigma X_M^* \cdot X_S (H_{LL} - H_{RL}) \dots (3)$$

10

$$\Sigma X_S^* \cdot 2E_L = \Sigma X_S^* \cdot X_M (H_{LL} + H_{RL}) + \Sigma X_S^* \cdot X_S (H_{LL} - H_{RL}) \dots (4)$$

are respectively obtained. Similarly, when both sides of the equation (2) are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$ ,

$$\Sigma X_M^* \cdot 2E_R = \Sigma X_M^* \cdot X_M (H_{LR} + H_{RR}) + \Sigma X_M^* \cdot X_S (H_{LR} - H_{RR}) \dots (5)$$

15

$$\Sigma X_S^* \cdot 2E_R = \Sigma X_S^* \cdot X_M (H_{LR} + H_{RR}) + \Sigma X_S^* \cdot X_S (H_{LR} - H_{RR}) \dots (6)$$

are respectively obtained.

[0070]

In the equations (3) to (6), since  $X_M$  and  $X_S$  are approximately uncorrelated with each other, such a term having  $X_M^* \cdot X_S$  or  $X_S^* \cdot X_M$  becomes approximately zero when ensemble-averaged. Further,

20

$$X_M^* \cdot X_M = |X_M|^2$$

$$X_S^* \cdot X_S = |X_S|^2$$

hence, the equations (3) to (6) respectively become

25

$$\Sigma X_M^* \cdot 2E_L = \Sigma |X_M|^2 (H_{LL} + H_{RL}) \dots (3')$$

$$\Sigma X_S^* \cdot 2E_L = \Sigma |X_S|^2 (H_{LL} - H_{RL}) \dots (4')$$

$$\Sigma X_M^* \cdot 2E_R = \Sigma |X_M|^2 (H_{LR} + H_{RR}) \dots (5')$$

$$\Sigma X_S^* \cdot 2E_R = \Sigma |X_S|^2 (H_{LR} - H_{RR}) \dots (6') .$$

By transforming the equations (3') to (6'), the following composite transfer functions each in the form of combination of two transfer functions are respectively  
5 derived.

$$H_{LL} + H_{RL} = \Sigma X_M^* \cdot 2E_L / \Sigma |X_M|^2 \dots (3'')$$

$$H_{LL} - H_{RL} = \Sigma X_S^* \cdot 2E_L / \Sigma |X_S|^2 \dots (4'')$$

$$H_{LR} + H_{RR} = \Sigma X_M^* \cdot 2E_R / \Sigma |X_M|^2 \dots (5'')$$

$$H_{LR} - H_{RR} = \Sigma X_S^* \cdot 2E_R / \Sigma |X_S|^2 \dots (6'')$$

10 When the corresponding sides of the equations (3'') and (4'') are added together,

$$H_{LL} = (\Sigma X_M^* \cdot E_L / \Sigma |X_M|^2) + (\Sigma X_S^* \cdot E_L / \Sigma |X_S|^2) \dots (7) .$$

When subtraction is performed between the corresponding sides of the equations (3'') and (4''),

15 
$$H_{RL} = (\Sigma X_M^* \cdot E_L / \Sigma |X_M|^2) - (\Sigma X_S^* \cdot E_L / \Sigma |X_S|^2) \dots (8) .$$

When the corresponding sides of the equations (5'') and (6'') are added together,

$$H_{LR} = (\Sigma X_M^* \cdot E_R / \Sigma |X_M|^2) + (\Sigma X_S^* \cdot E_R / \Sigma |X_S|^2) \dots (9) .$$

20 When subtraction is performed between the corresponding sides of the equations (5'') and (6''),

$$H_{RR} = (\Sigma X_M^* \cdot E_R / \Sigma |X_M|^2) - (\Sigma X_S^* \cdot E_R / \Sigma |X_S|^2) \dots (10) .$$

Hence, transfer functions  $H_{LL}$ ,  $H_{RL}$ ,  $H_{LR}$  and  $H_{RR}$  are respectively derived. Impulse responses  $h_{LL}$ ,  $h_{RL}$ ,  $h_{LR}$  and  $h_{RR}$  obtained by applying the inverse Fourier transformation to  
25 these derived transfer functions are the filter characteristics to be set to the filter means 40-1, 40-2, 40-3 and 40-4, respectively. Therefore, the transfer function

calculating means 58 derives the respective transfer functions  $H_{LL}$ ,  $H_{RL}$ ,  $H_{LR}$  and  $H_{RR}$  from the equations (7) to (10) based on the sum signal  $x_M$ , the difference signal  $x_S$ , and the output signals  $e_L$  and  $e_R$  of the subtracters 48 and 50 that  
5 are inputted, derives the impulse responses  $h_{LL}$ ,  $h_{RL}$ ,  $h_{LR}$  and  $h_{RR}$  by applying the inverse Fourier transformation to those derived transfer functions, sets the derived impulse responses to the filter means 40-1, 40-2, 40-3 and 40-4 as  $\hat{h}_{LL}$ ,  $\hat{h}_{RL}$ ,  $\hat{h}_{LR}$  and  $\hat{h}_{RR}$ , respectively, and further, updates  
10 the impulse responses by repeating this calculation per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

[0071]

When the foregoing impulse response updating  
15 technique is explained using estimated error parameters, it becomes as follows. Output signals (collected audio signals)  $y_L$  and  $y_R$  of the microphones MC(L) and MC(R), assuming that frequency-axis expressions of  $x_L$ ,  $x_R$ ,  $y_L$  and  $y_R$  are respectively given as  $X_L$ ,  $X_R$ ,  $Y_L$  and  $Y_R$ , become

$$20 \quad Y_L = (X_M + X_S) H_{LL} / 2 + (X_M - X_S) H_{RL} / 2$$

$$Y_R = (X_M + X_S) H_{LR} / 2 + (X_M - X_S) H_{RR} / 2.$$

The signal  $E_L$  outputted from the subtracter 48 becomes

$$\begin{aligned} E_L &= Y_L - (X_L \cdot \hat{H}_{LL} + X_R \cdot \hat{H}_{RL}) \\ 25 \quad &= \{ (X_M + X_S) H_{LL} / 2 + (X_M - X_S) H_{RL} / 2 \} - \{ (X_M + X_S) \hat{H}_{LL} / 2 \\ &\quad + (X_M - X_S) \hat{H}_{RL} / 2 \} \end{aligned}$$

hence

$$2E_L = X_M(H_{LL}+H_{RL}-H_{LL}^{\wedge}-H_{RL}^{\wedge}) + X_S(H_{LL}-H_{RL}-H_{LL}^{\wedge}+H_{RL}^{\wedge}) \dots (11).$$

The signal  $E_R$  outputted from the subtracter 50 becomes

$$\begin{aligned} E_R &= Y_R - (X_L \cdot H_{LR}^{\wedge} + X_R \cdot H_{RR}^{\wedge}) \\ 5 \quad &= \{ (X_M+X_S)H_{LR}/2 + (X_M-X_S)H_{RR}/2 \} - \{ (X_M+X_S)H_{LR}^{\wedge}/2 \\ &\quad + (X_M-X_S)H_{RR}^{\wedge}/2 \} \end{aligned}$$

hence

$$2E_R = X_M(H_{LR}+H_{RR}-H_{LR}^{\wedge}-H_{RR}^{\wedge}) + X_S(H_{LR}-H_{RR}-H_{LR}^{\wedge}+H_{RR}^{\wedge}) \dots (12).$$

[0072]

10 When the estimated errors are given as

$$\Delta H_{LL} = H_{LL} - H_{LL}^{\wedge}$$

$$\Delta H_{RL} = H_{RL} - H_{RL}^{\wedge}$$

$$\Delta H_{LR} = H_{LR} - H_{LR}^{\wedge}$$

$$\Delta H_{RR} = H_{RR} - H_{RR}^{\wedge}$$

15 the equations (11) and (12) become

$$2E_L = X_M(\Delta H_{LL}+\Delta H_{RL}) + X_S(\Delta H_{LL}-\Delta H_{RL}) \dots (11')$$

$$2E_R = X_M(\Delta H_{LR}+\Delta H_{RR}) + X_S(\Delta H_{LR}-\Delta H_{RR}) \dots (12').$$

When both sides of the equation (11') are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$  (i.e. deriving cross spectra) and ensemble-averaged,

20

$$\Sigma X_M^* \cdot 2E_L = \Sigma X_M^* \cdot X_M(\Delta H_{LL}+\Delta H_{RL}) + \Sigma X_M^* \cdot X_S(\Delta H_{LL}-\Delta H_{RL}) \dots (13)$$

$$\Sigma X_S^* \cdot 2E_L = \Sigma X_S^* \cdot X_M(\Delta H_{LL}+\Delta H_{RL}) + \Sigma X_S^* \cdot X_S(\Delta H_{LL}-\Delta H_{RL}) \dots (14)$$

25 are respectively obtained. Similarly, when both sides of the equation (12') are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$ ,

$$\Sigma X_M^* \cdot 2E_R = \Sigma X_M^* \cdot X_M (\Delta H_{LR} + \Delta H_{RR}) + \Sigma X_M^* \cdot X_S (\Delta H_{LR} - \Delta H_{RR}) \dots (15)$$

$$\Sigma X_S^* \cdot 2E_R = \Sigma X_S^* \cdot X_M (\Delta H_{LR} + \Delta H_{RR}) + \Sigma X_S^* \cdot X_S (\Delta H_{LR} - \Delta H_{RR}) \dots (16)$$

5 are respectively obtained.

[0073]

In the equations (13) to (16), since  $X_M$  and  $X_S$  are approximately uncorrelated with each other, such a term having  $X_M^* \cdot X_S$  or  $X_S^* \cdot X_M$  becomes approximately zero when  
10 ensemble-averaged. Further,

$$X_M^* \cdot X_M = |X_M|^2$$

$$X_S^* \cdot X_S = |X_S|^2$$

hence, the equations (13) to (16) respectively become

$$\Sigma X_M^* \cdot 2E_L = \Sigma |X_M|^2 (\Delta H_{LL} + \Delta H_{RL}) \dots (13')$$

$$\Sigma X_S^* \cdot 2E_L = \Sigma |X_S|^2 (\Delta H_{LL} - \Delta H_{RL}) \dots (14')$$

$$\Sigma X_M^* \cdot 2E_R = \Sigma |X_M|^2 (\Delta H_{LR} + \Delta H_{RR}) \dots (15')$$

$$\Sigma X_S^* \cdot 2E_R = \Sigma |X_S|^2 (\Delta H_{LR} - \Delta H_{RR}) \dots (16')$$

From the equations (13') to (16'),

$$\Delta H_{LL} = \Sigma X_M^* \cdot E_L / \Sigma |X_M|^2 + \Sigma X_S^* \cdot E_L / \Sigma |X_S|^2 \dots (17)$$

$$\Delta H_{RL} = \Sigma X_M^* \cdot E_L / \Sigma |X_M|^2 - \Sigma X_S^* \cdot E_L / \Sigma |X_S|^2 \dots (18)$$

$$\Delta H_{LR} = \Sigma X_M^* \cdot E_R / \Sigma |X_M|^2 + \Sigma X_S^* \cdot E_R / \Sigma |X_S|^2 \dots (19)$$

$$\Delta H_{RR} = \Sigma X_M^* \cdot E_R / \Sigma |X_M|^2 - \Sigma X_S^* \cdot E_R / \Sigma |X_S|^2 \dots (20)$$

are respectively derived.

[0074]

25 Using the estimated errors  $\Delta H_{LL}$ ,  $\Delta H_{RL}$ ,  $\Delta H_{LR}$  and  $\Delta H_{RR}$  derived from the equations (17) to (20), the filter characteristics of the filter means 40-1, 40-2, 40-3 and 40-4

are updated per suitably determined prescribed time period (e.g. time period of performing ensemble averaging). For example, assuming that impulse responses  $h_{LL}$ ,  $h_{RL}$ ,  $h_{LR}$  and  $h_{RR}$  after K-th updating are given as  $h_{LL}(k)$ ,  $h_{RL}(k)$ ,  $h_{LR}(k)$  and  $h_{RR}(k)$ , using impulse responses  $\Delta h_{LL}$ ,  $\Delta h_{RL}$ ,  $\Delta h_{LR}$  and  $\Delta h_{RR}$  corresponding to the derived estimated errors  $\Delta H_{LL}$ ,  $\Delta H_{RL}$ ,  $\Delta H_{LR}$  and  $\Delta H_{RR}$ ,

$$h_{LL}(k+1) = h_{LL}(k) + \alpha \Delta h_{LL} \dots (21)$$

$$h_{RL}(k+1) = h_{RL}(k) + \alpha \Delta h_{RL} \dots (22)$$

$$h_{LR}(k+1) = h_{LR}(k) + \alpha \Delta h_{LR} \dots (23)$$

$$h_{RR}(k+1) = h_{RR}(k) + \alpha \Delta h_{RR} \dots (24)$$

where  $\alpha$  is a suitably set convergence coefficient.

Using these updating equations, (k+1)th impulse responses  $h_{LL}(k+1)$ ,  $h_{RL}(k+1)$ ,  $h_{LR}(k+1)$  and  $h_{RR}(k+1)$  are derived and set to the filter means 40-1, 40-2, 40-3 and 40-4, respectively, which is repeated per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

[0075]

The results of carrying out simulations using the signals  $x_L$  and  $x_R$  or the sum and difference signals  $x_M$  and  $x_S$  as reference signals with respect to the stereo echo canceller 16, 24 of Fig. 1, are shown in Figs. 3 to 8 for each of the audio transfer systems. As the signals  $x_L$  and  $x_R$ , stereo audio signals based on a human voice were used. In Figs. 3 to 8, the axis of abscissas represents the number of blocks (one block represents a time period of performing

ensemble averaging and is set to about 2.3 seconds in the simulations), and the filter characteristics are updated per block. The filter characteristics are not set in the first block, the initial setting is executed in the second block, then updating is carried out per block. The axis of ordinates represents the estimated error (dB) of the transfer function, and the initial state where the filter characteristics are not set is defined as 0dB. Table 1 shows the conditions of the respective simulations of Figs. 3 to 8.

[Table 1]

Figure Number	Reference Signal	Double Talk	Change in Transfer System
FIG. 3	$X_L, X_R$	NO	NO
FIG. 4	$X_M, X_S$	NO	NO
FIG. 5	$X_L, X_R$	YES	NO
FIG. 6	$X_M, X_S$	YES	NO
FIG. 7	$X_L, X_R$	YES	YES
FIG. 8	$X_M, X_S$	YES	YES

[0076]

In Table 1, "Double Talk" represents the state where sounds reproduced from the loudspeakers SP(L) and SP(R) as well as a voice uttered by a person present in that room are simultaneously collected by the microphones MC(L) and MC(R). Since the teleconferencing system is used in general in the state where the double talk occurs, it is required that the

sufficient echo cancel performance can be achieved even when the double talk occurs. On the other hand, in Table 1, "Change in Transfer System" represents changing the transfer functions while the estimated errors are converging, supposing a case where the microphone positions are moved, or the like. As an operation of the echo canceller, it is required that even if the estimated error temporarily increases due to change in transfer function, it again go toward convergence.

10 [0077]

The simulation results of Figs. 3 to 8 are considered. Comparing Figs. 3 and 4, when the signals  $x_L$  and  $x_R$  are used as reference signals (Fig. 3), because a correlation between the signals  $x_L$  and  $x_R$  is high, estimated errors only drop to about -15 to -25dB at most in the 20th block (about 45 seconds) from the start of the operation. In contrast, when the signals  $x_M$  and  $x_S$  are used as reference signals (Fig. 4), because a correlation between the signals  $x_M$  and  $x_S$  is low, estimated errors drop to about -35 to -45dB in the 20th block from the start of the operation, and thus it is seen that the sufficient echo cancel performance can be achieved. Comparing Figs. 5 and 6 showing cases where the double talk exists, when the signals  $x_L$  and  $x_R$  are used as reference signals (Fig. 5), estimated errors only drop to about -10 to -20dB at most in the 20th block from the start of the operation. In contrast, when the signals  $x_M$  and  $x_S$  are used as reference signals (Fig. 6), estimated errors drop

to about -23 to -30dB, and thus it is seen that even if the double talk exists, the sufficient echo cancel performance can be achieved. This is because, since a voice uttered by a person present in a room is uncorrelated with sounds

5 reproduced from the loudspeakers  $SP(L)$  and  $SP(R)$ , when the transfer function calculating means 58 calculates transfer functions of the respective systems, components of the voice uttered by the person in the room are canceled through the foregoing cross-spectrum calculation and ensemble-average

10 calculation, so that the transfer functions of the respective systems can be derived with no influence of the double talk.

Comparing Figs. 7 and 8 showing cases where the double talk and the change of the transfer systems are present, when the signals  $x_L$  and  $x_R$  are used as reference signals (Fig. 7),

15 convergence of estimated errors is poor after giving a change to the transfer systems in the 11th block so that the estimated errors only drop to about -5 to -15dB at most in the 20th block. In contrast, when the signals  $x_M$  and  $x_S$  are used as reference signals (Fig. 8), convergence of estimated

20 errors is excellent even after the change is given to the transfer systems in the 11th block so that the estimated errors drop to about -17 to -26dB in the 20th block, and thus it is seen that even if the double talk and the change of the transfer systems exist, the sufficient echo cancel

25 performance can be achieved.

[0078]

Fig. 9 shows another structural example in the

stereo echo canceller 16, 24 of Fig. 2, wherein  
sum/difference signal producing means is arranged on  
transmission lines to loudspeakers. The same symbols are  
used with respect to those portions common to the foregoing  
5 structure of Fig. 1. Left/right two-channel stereo signals  
 $x_L$  and  $x_R$  transmitted from the spot on the counterpart side  
and inputted into line input ends LI(L) and LI(R) are  
inputted into sum/difference signal producing means 52. The  
sum/difference signal producing means 52 performs addition of  
10 the stereo signals  $x_L$  and  $x_R$  using an adder 54 so as to  
produce a sum signal  $x_M (= x_L + x_R)$ , while performs subtraction  
thereof using a subtracter 56 so as to produce a difference  
signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ . Stereo audio  
signal demodulating means 62 performs addition of the sum and  
15 difference signals  $x_M$  and  $x_S$  using an adder 64, and further,  
gives thereto a coefficient  $1/2$  using a coefficient  
multiplier 66 to recover the original signal  $x_L$ , while  
performs subtraction of the sum and difference signals  $x_M$  and  
 $x_S$  using a subtracter 68, and further, gives thereto a  
20 coefficient  $1/2$  using a coefficient multiplier 70 to recover  
the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are  
outputted from sound output ends SO(L) and SO(R) and  
reproduced at loudspeakers SP(L) and SP(R), respectively.

[0079]

25           Transfer function calculating means 58 implements a  
cross-spectrum calculation between the sum signal  $x_M$  and the  
difference signal  $x_S$  produced by the sum/difference signal

producing means 52 and signals  $e_L$  and  $e_R$  outputted from subtracters 48 and 50 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics of filter means 40-1 to 40-4. Operations thereof are the same as those described with respect to the structure of Fig. 1. Operations of the other portions are also the same as those described with respect to the structure of Fig. 1.

[0080]

Fig. 10 shows another structural example in the stereo echo canceller 16, 24 of Fig. 2, wherein transmission is implemented between the spots A and B of Fig. 2 in the signal form of the sum signal  $x_M$  and the difference signal  $x_S$ , instead of the signal form of the stereo signals  $x_L$  and  $x_R$ . The same symbols are used with respect to those portions common to the foregoing structure of Fig. 1 or 9. A sum signal  $x_M (= x_L + x_R)$  and a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L) \}$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are inputted into stereo audio signal demodulating means 62. The stereo audio signal demodulating means 62 performs addition of the sum and difference signals  $x_M$  and  $x_S$  using an adder 64, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 66 to recover the original signal  $x_L$ , while performs subtraction of the sum and difference signals  $x_M$  and  $x_S$  using a subtracter 68, and further, gives thereto a coefficient  $1/2$  using a coefficient

multiplier 70 to recover the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are outputted from sound output ends SO(L) and SO(R) and reproduced at loudspeakers SP(L) and SP(R), respectively.

5 [0081]

Transfer function calculating means 58 implements a cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  inputted from the line input ends LI(L) and LI(R) and signals  $e_L$  and  $e_R$  outputted from subtracters 48 and 50 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics of filter means 40-1 to 40-4. Operations thereof are the same as those described with respect to the structure of Fig. 1 or 9. Sum/difference signal producing means 72 performs addition, using an adder 73, of the signals  $e_L$  and  $e_R$  outputted from the subtracters 48 and 50 so as to produce a sum signal  $e_M (= e_L + e_R)$ , while performs subtraction thereof using a subtracter 75 so as to produce a difference signal  $e_S (= e_L - e_R$  (or it may also be  $e_R - e_L$ )), then sends them toward the spot on the counterpart side. Operations of the other portions are the same as those described with respect to the structure of Fig. 1 or 9.

[0082]

Fig. 11 shows another structural example in the stereo echo canceller 16, 24 of Fig. 2. The same symbols are used with respect to those portions common to the foregoing structure of Fig. 1, 9 or 10. A sum signal  $x_M (= x_L + x_R)$  and a

difference signal  $x_s \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$   
transmitted from the spot on the counterpart side and  
inputted into line input ends LI(L) and LI(R) are inputted  
into stereo audio signal demodulating means 62. The stereo  
5 audio signal demodulating means 62 performs addition of the  
sum and difference signals  $x_M$  and  $x_s$  using an adder 64, and  
further, gives thereto a coefficient  $1/2$  using a coefficient  
multiplier 66 to recover the original signal  $x_L$ , while  
performs subtraction of the sum and difference signals  $x_M$  and  
10  $x_s$  using a subtracter 68, and further, gives thereto a  
coefficient  $1/2$  using a coefficient multiplier 70 to recover  
the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are  
outputted from sound output ends SO(L) and SO(R) and  
reproduced at loudspeakers SP(L) and SP(R), respectively.

15 [0083]

Sum/difference signal producing means 52 performs  
addition, using an adder 54, of the stereo signals  $x_L$  and  $x_R$   
recovered by the stereo audio signal demodulating means 62 so  
as to produce a sum signal  $x_M (= x_L + x_R)$ , while performs  
20 subtraction thereof using a subtracter 56 so as to produce a  
difference signal  $x_s \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ .  
Transfer function calculating means 58 implements a cross-  
spectrum calculation between the sum signal  $x_M$  and the  
difference signal  $x_s$  produced by the sum/difference signal  
25 producing means 52 and signals  $e_L$  and  $e_R$  outputted from  
subtracters 48 and 50 and, based on this cross-spectrum  
calculation, performs setting and updating of filter

characteristics of filter means 40-1 to 40-4. Operations thereof are the same as those described with respect to the structure of Fig. 1, 9 or 10. Sum/difference signal producing means 72 performs addition, using an adder 73, of the signals  $e_L$  and  $e_R$  outputted from the subtracters 48 and 50 so as to produce a sum signal  $e_M (= e_L + e_R)$ , while performs subtraction thereof using a subtracter 75 so as to produce a difference signal  $e_S \{= e_L - e_R \text{ (or it may also be } e_R - e_L)\}$ , then sends them toward the spot on the counterpart side.

Operations of the other portions are the same as those described with respect to the structure of Fig. 1, 9 or 10.

[0084]

Fig. 12 shows a structural example in the stereo echo canceller 16, 24. Left/right two-channel stereo signals  $x_L$  and  $x_R$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are outputted from sound output ends SO(L) and SO(R) as they are (i.e. not through sum/difference signal producing means 152), and reproduced at loudspeakers SP(L) and SP(R), respectively.

[0085]

The sum/difference signal producing means 152 performs addition, using an adder 154, of the left/right two-channel stereo signals  $x_L$  and  $x_R$  inputted into the line input ends LI(L) and LI(R) so as to produce a sum signal  $x_M (= x_L + x_R)$ , while performs subtraction thereof using a subtracter 156 so as to produce a difference signal  $x_S \{= x_L - x_R \text{ (or } x_R - x_L)\}$ .

[0086]

Filer means 140-1 to 140-4 are formed by, for example, FIR filters. These filter means 140-1 to 140-4 are each set with an impulse response corresponding to a  
5 composite transfer function in the form of combination of transfer functions of suitable two systems among transfer functions  $H_{LL}$ ,  $H_{LR}$ ,  $H_{RL}$  and  $H_{RR}$  of four audio transfer systems between the loudspeakers SP(L) and SP(R) and microphones MC(L) and MC(R), respectively, and perform a convolution  
10 calculation of the sum and difference signals (low-correlation composite signals) using such impulse responses, thereby producing echo cancel signals EC1 to EC4, respectively.

[0087]

15 An adder 144 performs a calculation of  $EC1+EC3$ . An adder 146 performs a calculation of  $EC2+EC4$ . A subtracter 148 subtracts an echo cancel signal  $EC1+EC3$  from a collected audio signal of the microphone MC(L) inputted from a sound input end SI(L), thereby to perform echo cancellation. A  
20 subtracter 150 subtracts an echo cancel signal  $EC2+EC4$  from a collected audio signal of the microphone MC(R) inputted from a sound input end SI(R), thereby to perform echo cancellation. These echo-canceled signals of the respective left and right channels are outputted from line output ends LO(L) and LO(R),  
25 respectively, and transmitted toward the spot on the counterpart side.

[0088]

Transfer function calculating means 158 implements a cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal producing means 152 and signals  $e_L$  and  $e_R$  outputted from the subtracters 148 and 150 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics (impulse responses) of the filter means 140-1 to 140-4. Specifically, upon starting the system, the filter characteristics of the filter means 140-1 to 140-4 are not set, i.e. coefficients are all set to zero, so that the echo cancel signals EC1 to EC4 are zero, and thus the collected audio signals of the microphones MC(L) and MC(R) themselves are outputted from the subtracters 148 and 150. Therefore, at this time, the transfer function calculating means 158 performs the cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal producing means 152 and the collected audio signals  $e_L$  and  $e_R$  of the microphones MC(L) and MC(R) outputted from the subtracters 148 and 150 and, based on this cross-spectrum calculation, derives a plurality of composite transfer functions each in the form of combination of transfer functions of suitable two systems among transfer functions  $H_{LL}$ ,  $H_{LR}$ ,  $H_{RL}$  and  $H_{RR}$  of four audio transfer systems between the loudspeakers SP(L) and SP(R) and the microphones MC(L) and MC(R), respectively, and implements initial setting of the filter characteristics of the filter means 140-1 to 140-4 to values corresponding to such composite transfer

functions. After the initial setting, since the echo cancel signals are produced by the filter means 140-1 to 140-4, the echo cancel error signals  $e_L$  and  $e_R$  corresponding to difference signals between the collected audio signals of the microphones MC(L) and MC(R) and the echo cancel signals EC1 to EC4 are outputted from the subtracters 148 and 150. Therefore, at this time, the transfer function calculating means 158 performs the cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal producing means 152 and the echo cancel error signals  $e_L$  and  $e_R$  outputted from the subtracters 148 and 150 and, based on this cross-spectrum calculation, derives estimated errors of the foregoing composite transfer functions, respectively, and updates the filter characteristics of the filter means 140-1 to 140-4 to values that cancel such estimated errors, respectively. By repeating this updating operation per prescribed time period, the echo cancel error can be converged to a minimum value. Further, even if the transfer functions change due to movement of the microphone positions or the like, the echo cancel error can be converged to a minimum value by sequentially updating the filter characteristics of the filter means 140-1 to 140-4 depending thereon.

[0089]

Correlation detecting means 160 detects a correlation between the sum signal  $x_M$  and the difference signal  $x_S$  based on a correlation value calculation or the

like, and stops updating of the foregoing filter characteristics when the correlation value is no less than a prescribed value. When the correlation value becomes lower than the prescribed value, updating of the foregoing filter characteristics is restarted.

[0090]

Herein, the filter characteristics (impulse responses) that are set to the filter means 140-1 to 140-4 by the transfer function calculating means 158 will be described. In the transfer function calculating means 158, the following calculation is performed.

(In case of Fixed Type Operation)

The signals  $x_L$  and  $x_R$  are

$$x_L = (x_M + x_S) / 2$$

$$x_R = (x_M - x_S) / 2$$

hence, output signals  $Y_L$  and  $Y_R$  of the microphones MC(L) and MC(R) become

$$\begin{aligned} Y_L &= (x_M + x_S) H_{LL} / 2 + (x_M - x_S) H_{RL} / 2 \\ &= x_M \{ (H_{LL} + H_{RL}) / 2 \} + x_S \{ (H_{LL} - H_{RL}) / 2 \} \dots (25) \end{aligned}$$

$$\begin{aligned} Y_R &= (x_M + x_S) H_{LR} / 2 + (x_M - x_S) H_{RR} / 2 \\ &= x_M \{ (H_{LR} + H_{RR}) / 2 \} + x_S \{ (H_{LR} - H_{RR}) / 2 \} \dots (26) . \end{aligned}$$

When the composite transfer functions are given as

$$H_{ML} = (H_{LL} + H_{RL}) / 2$$

$$H_{SL} = (H_{LL} - H_{RL}) / 2$$

$$H_{MR} = (H_{LR} + H_{RR}) / 2$$

$$H_{SR} = (H_{LR} - H_{RR}) / 2$$

the equations (25) and (26) respectively become

$$Y_L = X_M \cdot H_{ML} + X_S \cdot H_{SL} \dots (25')$$

$$Y_R = X_M \cdot H_{MR} + X_S \cdot H_{SR} \dots (26')$$

When both sides of the equations (25') and (26') are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$  and ensemble-averaged,

$$\Sigma X_M^* \cdot Y_L = \Sigma X_M^* \cdot X_M \cdot H_{ML} + \Sigma X_M^* \cdot X_S \cdot H_{SL} \dots (27)$$

$$\Sigma X_S^* \cdot Y_L = \Sigma X_S^* \cdot X_M \cdot H_{ML} + \Sigma X_S^* \cdot X_S \cdot H_{SL} \dots (28)$$

$$\Sigma X_M^* \cdot Y_R = \Sigma X_M^* \cdot X_M \cdot H_{MR} + \Sigma X_M^* \cdot X_S \cdot H_{SR} \dots (29)$$

$$\Sigma X_S^* \cdot Y_R = \Sigma X_S^* \cdot X_M \cdot H_{MR} + \Sigma X_S^* \cdot X_S \cdot H_{SR} \dots (30)$$

are respectively obtained.

[0091]

In the equations (27) to (30), since  $X_M$  and  $X_S$  are approximately uncorrelated with each other, such a term having  $X_M^* \cdot X_S$  or  $X_S^* \cdot X_M$  becomes approximately zero when ensemble-averaged. Further,

$$X_M^* \cdot X_M = |X_M|^2$$

$$X_S^* \cdot X_S = |X_S|^2$$

hence, the equations (27) to (30) respectively become

$$\Sigma X_M^* \cdot Y_L = \Sigma |X_M|^2 H_{ML} \dots (27')$$

$$\Sigma X_S^* \cdot Y_L = \Sigma |X_S|^2 H_{SL} \dots (28')$$

$$\Sigma X_M^* \cdot Y_R = \Sigma |X_M|^2 H_{MR} \dots (29')$$

$$\Sigma X_S^* \cdot Y_R = \Sigma |X_S|^2 H_{SR} \dots (30')$$

From the equations (27') to (30'),

$$H_{ML} = \Sigma X_M^* \cdot Y_L / \Sigma |X_M|^2 \dots (31)$$

$$H_{SL} = \Sigma X_S^* \cdot Y_L / \Sigma |X_S|^2 \dots (32)$$

$$H_{MR} = \Sigma X_M^* \cdot Y_R / \Sigma |X_M|^2 \dots (33)$$

$$H_{SR} = \Sigma X_S^* \cdot Y_R / \Sigma |X_S|^2 \dots (34)$$

are respectively derived.

[0092]

Impulse responses  $h_{ML}$ ,  $h_{SL}$ ,  $h_{MR}$  and  $h_{SR}$  obtained by applying the inverse Fourier transformation to these derived composite transfer functions  $H_{ML}$ ,  $H_{SL}$ ,  $H_{MR}$  and  $H_{SR}$  are the filter characteristics to be set to the filter means 140-1, 140-2, 140-3 and 140-4, respectively. Therefore, the transfer function calculating means 158 derives the respective composite transfer functions  $H_{ML}$ ,  $H_{SL}$ ,  $H_{MR}$  and  $H_{SR}$  from the equations (31) to (34) based on the sum signal  $x_M$ , the difference signal  $x_S$ , and output signals  $y_L$  and  $y_R$  of the microphones MC(L) and MC(R) that are inputted, derives the impulse responses  $h_{ML}$ ,  $h_{SL}$ ,  $h_{MR}$  and  $h_{SR}$  by applying the inverse Fourier transformation to those derived composite transfer functions, sets the derived impulse responses to the filter means 140-1, 140-2, 140-3 and 140-4, respectively, and further, updates the impulse responses by repeating this calculation per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

[0093]

(In case of Adaptive Type Operation)

Assuming that the filter characteristics set to the filter means 140-1, 140-2, 140-3 and 140-4 are given as  $\hat{H}_{ML}$ ,  $\hat{H}_{SL}$ ,  $\hat{H}_{MR}$  and  $\hat{H}_{SR}$  ( $\hat{h}_{ML}$ ,  $\hat{h}_{SL}$ ,  $\hat{h}_{MR}$  and  $\hat{h}_{SR}$  when expressed in terms of the impulse responses), the signals  $e_L$  and  $e_R$  outputted from the subtractors 148 and 150 become

$$E_L = Y_L - (X_M \cdot \hat{H}_{ML} + X_S \cdot \hat{H}_{SL})$$

$$= \{ (X_M + X_S) H_{LL} / 2 + (X_M - X_S) H_{RL} / 2 \} - (X_M \cdot H^{\wedge}_{ML} + X_S \cdot H^{\wedge}_{SL})$$

$$= X_M \{ (H_{LL} + H_{RL}) / 2 - H^{\wedge}_{ML} \} + X_S \{ (H_{LL} - H_{RL}) / 2 - H^{\wedge}_{SL} \} \dots (35)$$

$$E_R = Y_R - (X_M \cdot H^{\wedge}_{MR} + X_S \cdot H^{\wedge}_{SR})$$

$$= \{ (X_M + X_S) H_{LR} / 2 + (X_M - X_S) H_{RR} / 2 \} - (X_M \cdot H^{\wedge}_{MR} + X_S \cdot H^{\wedge}_{SR})$$

$$= X_M \{ (H_{LR} + H_{RR}) / 2 - H^{\wedge}_{MR} \} + X_S \{ (H_{LR} - H_{RR}) / 2 - H^{\wedge}_{SR} \} \dots (36)$$

When the composite transfer functions are given as

$$H_{ML} = (H_{LL} + H_{RL}) / 2$$

$$H_{SL} = (H_{LL} - H_{RL}) / 2$$

$$H_{MR} = (H_{LR} + H_{RR}) / 2$$

$$H_{SR} = (H_{LR} - H_{RR}) / 2$$

the equations (35) and (36) respectively become

$$E_L = X_M (H_{ML} - H^{\wedge}_{ML}) + X_S (H_{SL} - H^{\wedge}_{SL}) \dots (35')$$

$$E_R = X_M (H_{MR} - H^{\wedge}_{MR}) + X_S (H_{SR} - H^{\wedge}_{SR}) \dots (36')$$

When the estimated errors of the composite transfer

functions are given as

$$\Delta H_{ML} = H_{ML} - H^{\wedge}_{ML}$$

$$\Delta H_{SL} = H_{SL} - H^{\wedge}_{SL}$$

$$\Delta H_{MR} = H_{MR} - H^{\wedge}_{MR}$$

$$\Delta H_{SR} = H_{SR} - H^{\wedge}_{SR}$$

the equations (35') and (36') respectively become

$$E_L = X_M \cdot \Delta H_{ML} + X_S \cdot \Delta H_{SL} \dots (35'')$$

$$E_R = X_M \cdot \Delta H_{MR} + X_S \cdot \Delta H_{SR} \dots (36'')$$

When both sides of the equations (35'') and (36'') are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$  and

ensemble-averaged,

$$\Sigma X_M^* \cdot E_L = \Sigma X_M^* \cdot X_M \cdot \Delta H_{ML} + \Sigma X_M^* \cdot X_S \cdot \Delta H_{SL} \dots (37)$$

$$\Sigma X_S^* \cdot E_L = \Sigma X_S^* \cdot X_M \cdot \Delta H_{ML} + \Sigma X_S^* \cdot X_S \cdot \Delta H_{SL} \dots (38)$$

$$\Sigma X_M^* \cdot E_R = \Sigma X_M^* \cdot X_M \cdot \Delta H_{MR} + \Sigma X_M^* \cdot X_S \cdot \Delta H_{SR} \dots (39)$$

$$\Sigma X_S^* \cdot E_R = \Sigma X_S^* \cdot X_M \cdot \Delta H_{MR} + \Sigma X_S^* \cdot X_S \cdot \Delta H_{SR} \dots (40)$$

are respectively obtained.

[0094]

5 In the equations (37) to (40), since  $X_M$  and  $X_S$  are approximately uncorrelated with each other, such a term having  $X_M^* \cdot X_S$  or  $X_S^* \cdot X_M$  becomes approximately zero when ensemble-averaged. Further,

$$X_M^* \cdot X_M = |X_M|^2$$

10  $X_S^* \cdot X_S = |X_S|^2$

hence, the equations (37) to (40) respectively become

$$\Sigma X_M^* \cdot E_L = \Sigma |X_M|^2 \Delta H_{ML} \dots (37')$$

$$\Sigma X_S^* \cdot E_L = \Sigma |X_S|^2 \Delta H_{SL} \dots (38')$$

$$\Sigma X_M^* \cdot E_R = \Sigma |X_M|^2 \Delta H_{MR} \dots (39')$$

15  $\Sigma X_S^* \cdot E_R = \Sigma |X_S|^2 \Delta H_{SR} \dots (40')$

From the equations (37') to (40'),

$$\Delta H_{ML} = \Sigma X_M^* \cdot E_L / \Sigma |X_M|^2 \dots (41)$$

$$\Delta H_{SL} = \Sigma X_S^* \cdot E_L / \Sigma |X_S|^2 \dots (42)$$

$$\Delta H_{MR} = \Sigma X_M^* \cdot E_R / \Sigma |X_M|^2 \dots (43)$$

20  $\Delta H_{SR} = \Sigma X_S^* \cdot E_R / \Sigma |X_S|^2 \dots (44)$

are respectively derived.

[0095]

Using the estimated errors  $\Delta H_{ML}$ ,  $\Delta H_{SL}$ ,  $\Delta H_{MR}$  and  $\Delta H_{SR}$  derived from the equations (40) to (44), the filter  
25 characteristics of the filter means 140-1, 140-2, 140-3 and 140-4 are updated per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

For example, assuming that impulse responses  $h_{ML}$ ,  $h_{SL}$ ,  $h_{MR}$  and  $h_{SR}$  after K-th updating are given as  $h_{ML}(k)$ ,  $h_{SL}(k)$ ,  $h_{MR}(k)$  and  $h_{SR}(k)$ , using impulse responses  $\Delta h_{ML}$ ,  $\Delta h_{SL}$ ,  $\Delta h_{MR}$  and  $\Delta h_{SR}$  corresponding to the derived estimated errors  $\Delta H_{ML}$ ,  $\Delta H_{SL}$ ,  $\Delta H_{MR}$  and  $\Delta H_{SR}$ ,

$$h_{ML}(k+1) = h_{ML}(k) + \alpha \Delta h_{ML} \dots (45)$$

$$h_{SL}(k+1) = h_{SL}(k) + \alpha \Delta h_{SL} \dots (46)$$

$$h_{MR}(k+1) = h_{MR}(k) + \alpha \Delta h_{MR} \dots (47)$$

$$h_{SR}(k+1) = h_{SR}(k) + \alpha \Delta h_{SR} \dots (48).$$

Using these updating equations, (k+1)th impulse responses  $h_{ML}(k+1)$ ,  $h_{SL}(k+1)$ ,  $h_{MR}(k+1)$  and  $h_{SR}(k+1)$  are derived and set to the filter means 140-1, 140-2, 140-3 and 140-4, respectively, which is repeated per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

[0096]

Fig. 13 shows another structural example in the stereo echo canceller 16, 24 of Fig. 2, wherein sum/difference signal producing means is arranged on transmission lines to loudspeakers. The same symbols are used with respect to those portions common to the foregoing structure of Fig. 12. Left/right two-channel stereo signals  $x_L$  and  $x_R$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are inputted into sum/difference signal producing means 152. The sum/difference signal producing means 152 performs addition of the stereo signals  $x_L$  and  $x_R$  using an adder 154 so as to

produce a sum signal  $x_M (= x_L + x_R)$ , while performs subtraction thereof using a subtracter 156 so as to produce a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ . Stereo audio signal demodulating means 162 performs addition of the sum and difference signals  $x_M$  and  $x_S$  using an adder 164, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 166 to recover the original signal  $x_L$ , while performs subtraction of the sum and difference signals  $x_M$  and  $x_S$  using a subtracter 168, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 170 to recover the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are outputted from sound output ends  $SO(L)$  and  $SO(R)$  and reproduced at loudspeakers  $SP(L)$  and  $SP(R)$ , respectively.

[0097]

Transfer function calculating means 158 implements a cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal producing means 152 and signals  $e_L$  and  $e_R$  outputted from subtracters 148 and 150 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics of filter means 140-1 to 140-4. Operations thereof are the same as those described with respect to the structure of Fig. 12. Operations of the other portions are also the same as those described with respect to the structure of Fig. 12.

[0098]

Fig. 14 shows another structural example in the

stereo echo canceller 16, 24 of Fig. 2, wherein transmission is implemented between the spots A and B of Fig. 2 in the signal form of the sum signal  $x_M$  and the difference signal  $x_S$ , instead of the signal form of the stereo signals  $x_L$  and  $x_R$ .

5 The same symbols are used with respect to those portions common to the foregoing structure of Fig. 12 or 13. A sum signal  $x_M (= x_L + x_R)$  and a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and  
10 LI(R) are inputted into stereo audio signal demodulating means 162. The stereo audio signal demodulating means 162 performs addition of the sum and difference signals  $x_M$  and  $x_S$  using an adder 164, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 166 to recover the  
15 original signal  $x_L$ , while performs subtraction of the sum and difference signals  $x_M$  and  $x_S$  using a subtracter 168, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 170 to recover the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are outputted from sound output  
20 ends SO(L) and SO(R) and reproduced at loudspeakers SP(L) and SP(R), respectively.

[0099]

Transfer function calculating means 158 implements a cross-spectrum calculation between the sum signal  $x_M$  and the  
25 difference signal  $x_S$  inputted from the line input ends LI(L) and LI(R) and signals  $e_L$  and  $e_R$  outputted from subtracters 148 and 150 and, based on this cross-spectrum calculation,

performs setting and updating of filter characteristics of filter means 140-1 to 140-4. Operations thereof are the same as those described with respect to the structure of Fig. 12 or 13. Sum/difference signal producing means 172 performs  
5 addition, using an adder 173, of the signals  $e_L$  and  $e_R$  outputted from the subtracters 148 and 150 so as to produce a sum signal  $e_M (= e_L + e_R)$ , while performs subtraction thereof using a subtracter 175 so as to produce a difference signal  $e_S \{= e_L - e_R \text{ (or it may also be } e_R - e_L)\}$ , then sends them toward  
10 the spot on the counterpart side. Operations of the other portions are the same as those described with respect to the structure of Fig. 12 or 13.

[0100]

Fig. 15 shows another structural example in the  
15 stereo echo canceller 16, 24 of Fig. 2. The same symbols are used with respect to those portions common to the foregoing structure of Fig. 12, 13 or 14. A sum signal  $x_M (= x_L + x_R)$  and a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$  transmitted from the spot on the counterpart side and  
20 inputted into line input ends LI(L) and LI(R) are inputted into stereo audio signal demodulating means 162. The stereo audio signal demodulating means 162 performs addition of the sum and difference signals  $x_M$  and  $x_S$  using an adder 164, and further, gives thereto a coefficient  $1/2$  using a coefficient  
25 multiplier 166 to recover the original signal  $x_L$ , while performs subtraction of the sum and difference signals  $x_M$  and  $x_S$  using a subtracter 168, and further, gives thereto a

coefficient  $1/2$  using a coefficient multiplier 170 to recover the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are outputted from sound output ends SO(L) and SO(R) and reproduced at loudspeakers SP(L) and SP(R), respectively.

5 [0101]

Sum/difference signal producing means 152 performs addition, using an adder 154, of the stereo signals  $x_L$  and  $x_R$  recovered by the stereo audio signal demodulating means 162 so as to produce a sum signal  $x_M (= x_L + x_R)$ , while performs subtraction thereof using a subtracter 156 so as to produce a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ . Transfer function calculating means 158 implements a cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal producing means 152 and signals  $e_L$  and  $e_R$  outputted from subtracters 148 and 150 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics of filter means 140-1 to 140-4. Operations thereof are the same as those described with respect to the structure of Fig. 12, 13 or 14. Sum/difference signal producing means 172 performs addition, using an adder 173, of the signals  $e_L$  and  $e_R$  outputted from the subtracters 148 and 150 so as to produce a sum signal  $e_M (= e_L + e_R)$ , while performs subtraction thereof using a subtracter 175 so as to produce a difference signal  $e_S \{= e_L - e_R \text{ (or it may also be } e_R - e_L)\}$ , then sends them toward the spot on the counterpart side. Operations of the other portions are the same as those

described with respect to the structure of Fig. 12, 13 or 14.

[0102]

Fig. 16 shows another structural example in the stereo echo canceller 16, 24. A sum signal  $x_M (= x_L + x_R)$  and a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L) \}$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are inputted into stereo audio signal demodulating means 262. The stereo audio signal demodulating means 262 performs addition of the sum and difference signals  $x_M$  and  $x_S$  using an adder 264, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 266 to recover the original signal  $x_L$ , while performs subtraction of the sum and difference signals  $x_M$  and  $x_S$  using a subtracter 268, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 270 to recover the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are outputted from sound output ends SO(L) and SO(R) and reproduced at loudspeakers SP(L) and SP(R), respectively.

[0103]

Collected audio signals  $y_L$  and  $y_R$  of microphones MC(L) and MC(R) are inputted into sum/difference signal producing means 272. The sum/difference signal producing means 272 implements addition of the microphone collected audio signals  $y_L$  and  $y_R$  using an adder 273 so as to produce a sum signal  $y_M$ , while implements subtraction thereof using a subtracter 275 so as to produce a difference signal  $y_S$ .

[0104]

Filer means 240-1 to 240-4 are formed by, for example, FIR filters. These filter means 240-1 to 240-4 are each set with an impulse response corresponding to a composite transfer function in the form of combination of transfer functions of suitable two systems among transfer functions  $H_{LL}$ ,  $H_{LR}$ ,  $H_{RL}$  and  $H_{RR}$  of four audio transfer systems between the loudspeakers  $SP(L)$  and  $SP(R)$  and microphones  $MC(L)$  and  $MC(R)$ , respectively, and perform a convolution calculation of the left/right two-channel stereo signals  $x_L$  and  $x_R$  using such impulse responses, thereby producing echo cancel signals  $EC1$  to  $EC4$ , respectively.

[0105]

An adder 244 performs a calculation of  $EC1+EC3$ . An adder 246 performs a calculation of  $EC2+EC4$ . A subtracter 248 subtracts an echo cancel signal  $EC1+EC3$  from the sum signal  $y_M$ , thereby to perform echo cancellation. A subtracter 250 subtracts an echo cancel signal  $EC2+EC4$  from the difference signal  $y_S$ , thereby to perform echo cancellation. Signals  $e_M$  and  $e_S$  outputted from the subtracters 248 and 250 are outputted from line output ends  $LO(L)$  and  $LO(R)$ , respectively, and transmitted toward the spot on the counterpart side.

[0106]

Transfer function calculating means 258 implements a cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  inputted from the line input ends  $LI(L)$  and  $LI(R)$  and the signals  $e_M$  and  $e_S$  outputted from the

subtractors 248 and 250 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics (impulse responses) of the filter means 240-1 to 240-4. Specifically, upon starting the system, the filter characteristics of the filter means 240-1 to 240-4 are not set, i.e. coefficients are all set to zero, so that the echo cancel signals EC1 to EC4 are zero, and thus the sum signal  $y_M$  and the difference signal  $y_S$  outputted from the sum/difference signal producing means 272, as they are, are outputted from the subtractors 248 and 250. Therefore, at this time, the transfer function calculating means 258 performs the cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  inputted from the line input ends LI(L) and LI(R) and the sum signal  $y_M$  and the difference signal  $y_S$  outputted from the subtractors 248 and 250 and, based on this cross-spectrum calculation, derives a plurality of composite transfer functions each in the form of combination of transfer functions of suitable two systems among transfer functions  $H_{LL}$ ,  $H_{LR}$ ,  $H_{RL}$  and  $H_{RR}$  of four audio transfer systems between the loudspeakers SP(L) and SP(R) and the microphones MC(L) and MC(R), respectively, and implements initial setting of the filter characteristics of the filter means 240-1 to 240-4 to values corresponding to such composite transfer functions. After the initial setting, since the echo cancel signals are produced by the filter means 240-1 to 240-4, the echo cancel error signals  $e_M$  and  $e_S$  corresponding to difference signals between the sum signal  $y_M$

and the difference signal  $y_s$  outputted from the sum/difference signal producing means 272 and the echo cancel signals EC1 to EC4 are outputted from the subtracters 248 and 250. Therefore, at this time, the transfer function calculating means 258 performs the cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_s$  inputted from the line input ends LI(L) and LI(R) and the echo cancel error signals  $e_M$  and  $e_s$  outputted from the subtracters 248 and 250 and, based on this cross-spectrum calculation, derives estimated errors of the foregoing composite transfer functions, respectively, and updates the filter characteristics of the filter means 240-1 to 240-4 to values that cancel such estimated errors, respectively. By repeating this updating operation per prescribed time period, the echo cancel error can be converged to a minimum value. Further, even if the transfer functions change due to movement of the microphone positions or the like, the echo cancel error can be converged to a minimum value by sequentially updating the filter characteristics of the filter means 240-1 to 240-4 depending thereon.

[0107]

Correlation detecting means 260 detects a correlation between the sum signal  $x_M$  and the difference signal  $x_s$  based on a correlation value calculation or the like, and stops updating of the foregoing filter characteristics when the correlation value is no less than a prescribed value. When the correlation value becomes lower

than the prescribed value, updating of the foregoing filter characteristics is restarted.

[0108]

Herein, the filter characteristics (impulse responses) that are set to the filter means 240-1 to 240-4 by the transfer function calculating means 258 will be described. In the transfer function calculating means 258, the following calculation is performed.

(In case of Fixed Type Operation)

The sum and difference signals  $y_M$  and  $y_S$  of the output signals  $y_L$  and  $y_R$  of the microphones MC(L) and MC(R), assuming that frequency-axis expressions of  $y_M$  and  $y_S$  are respectively given as  $Y_M$  and  $Y_S$ , become

$$\begin{aligned} Y_M &= Y_L + Y_R \\ &= (X_L \cdot H_{LL} + X_R \cdot H_{RL}) + (X_L \cdot H_{LR} + X_R \cdot H_{RR}) \\ &= X_L (H_{LL} + H_{LR}) + X_R (H_{RL} + H_{RR}) \dots (49) \end{aligned}$$

$$\begin{aligned} Y_S &= Y_L - Y_R \\ &= (X_L \cdot H_{LL} + X_R \cdot H_{RL}) - (X_L \cdot H_{LR} + X_R \cdot H_{RR}) \\ &= X_L (H_{LL} - H_{LR}) + X_R (H_{RL} - H_{RR}) \dots (50) \end{aligned}$$

When the composite transfer functions are given as

$$H_{LM} = H_{LL} + H_{LR}$$

$$H_{RM} = H_{RL} + H_{RR}$$

$$H_{LS} = H_{LL} - H_{LR}$$

$$H_{RS} = H_{RL} - H_{RR}$$

then

$$X_L = (X_M + X_S) / 2$$

$$X_R = (X_M - X_S) / 2$$

hence, the equations (49) and (50) respectively become

$$\begin{aligned} Y_M &= (X_M + X_S) H_{LM} / 2 + (X_M - X_S) H_{RM} / 2 \\ &= X_M (H_{LM} + H_{RM}) / 2 + X_S (H_{LM} - H_{RM}) / 2 \\ Y_S &= (X_M + X_S) H_{LS} / 2 + (X_M - X_S) H_{RS} / 2 \\ &= X_M (H_{LS} + H_{RS}) / 2 + X_S (H_{LS} - H_{RS}) / 2 \end{aligned}$$

thus

$$2Y_M = X_M (H_{LM} + H_{RM}) + X_S (H_{LM} - H_{RM}) \dots (49')$$

$$2Y_S = X_M (H_{LS} + H_{RS}) + X_S (H_{LS} - H_{RS}) \dots (50')$$

When both sides of the equations (49') and (50') are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$  and ensemble-averaged,

$$\Sigma X_M^* \cdot Y_M = \Sigma X_M^* \cdot X_M (H_{LM} + H_{RM}) + \Sigma X_M^* \cdot X_S (H_{LM} - H_{RM}) \dots (51)$$

$$\Sigma X_S^* \cdot Y_M = \Sigma X_S^* \cdot X_M (H_{LM} + H_{RM}) + \Sigma X_S^* \cdot X_S (H_{LM} - H_{RM}) \dots (52)$$

$$\Sigma X_M^* \cdot Y_S = \Sigma X_M^* \cdot X_M (H_{LS} + H_{RS}) + \Sigma X_M^* \cdot X_S (H_{LS} - H_{RS}) \dots (53)$$

$$\Sigma X_S^* \cdot Y_S = \Sigma X_S^* \cdot X_M (H_{LS} + H_{RS}) + \Sigma X_S^* \cdot X_S (H_{LS} - H_{RS}) \dots (54)$$

are respectively obtained.

[0109]

In the equations (51) to (54), since  $X_M$  and  $X_S$  are approximately uncorrelated with each other, such a term having  $X_M^* \cdot X_S$  or  $X_S^* \cdot X_M$  becomes approximately zero when ensemble-averaged. Further,

$$X_M^* \cdot X_M = |X_M|^2$$

$$X_S^* \cdot X_S = |X_S|^2$$

hence, the equations (51) to (54) respectively become

$$\Sigma X_M^* \cdot Y_M = \Sigma |X_M|^2 (H_{LM} + H_{RM}) \dots (51')$$

$$\Sigma X_S^* \cdot Y_M = \Sigma |X_S|^2 (H_{LM} - H_{RM}) \dots (52')$$

$$\Sigma X_M^* \cdot Y_S = \Sigma |X_M|^2 (H_{LS} + H_{RS}) \dots (53')$$

$$\Sigma X_S^* \cdot Y_S = \Sigma |X_S|^2 (H_{LS} - H_{RS}) \dots (54')$$

By transforming the equations (51') to (54'), the following composite transfer functions are respectively derived.

$$5 \quad H_{LM} + H_{RM} = \Sigma X_M^* \cdot Y_M / \Sigma |X_M|^2 \dots (51'')$$

$$H_{LM} - H_{RM} = \Sigma X_S^* \cdot Y_M / \Sigma |X_S|^2 \dots (52'')$$

$$H_{LS} + H_{RS} = \Sigma X_M^* \cdot Y_S / \Sigma |X_M|^2 \dots (53'')$$

$$H_{LS} - H_{RS} = \Sigma X_S^* \cdot Y_S / \Sigma |X_S|^2 \dots (54'')$$

From the equations (51'') to (54''),

$$10 \quad H_{LM} = \Sigma X_M^* \cdot Y_M / \Sigma |X_M|^2 + \Sigma X_S^* \cdot Y_M / \Sigma |X_S|^2 \dots (55)$$

$$H_{RM} = \Sigma X_M^* \cdot Y_M / \Sigma |X_M|^2 - \Sigma X_S^* \cdot Y_M / \Sigma |X_S|^2 \dots (56)$$

$$H_{LS} = \Sigma X_M^* \cdot Y_S / \Sigma |X_M|^2 + \Sigma X_S^* \cdot Y_S / \Sigma |X_S|^2 \dots (57)$$

$$H_{RS} = \Sigma X_M^* \cdot Y_S / \Sigma |X_M|^2 - \Sigma X_S^* \cdot Y_S / \Sigma |X_S|^2 \dots (58)$$

are respectively derived.

$$15 \quad [0110]$$

Impulse responses  $h_{LM}$ ,  $h_{RM}$ ,  $h_{LS}$  and  $h_{RS}$  obtained by applying the inverse Fourier transformation to these derived composite transfer functions  $H_{LM}$ ,  $H_{RM}$ ,  $H_{LS}$  and  $H_{RS}$  are the filter characteristics to be set to the filter means 240-1,

20 240-2, 240-3 and 240-4, respectively. Therefore, the

transfer function calculating means 258 derives the

respective composite transfer functions  $H_{LM}$ ,  $H_{RM}$ ,  $H_{LS}$  and  $H_{RS}$  from the equations (55) to (58) based on the sum signal  $x_M$

and the difference signal  $x_S$  inputted into the line input

25 ends LI(L) and LI(R) and the sum signal  $y_M$  and the difference

signal  $y_S$  outputted from the sum/difference signal producing

means 272, derives the impulse responses  $h_{LM}$ ,  $h_{RM}$ ,  $h_{LS}$  and  $h_{RS}$

by applying the inverse Fourier transformation to those derived composite transfer functions, sets the derived impulse responses to the filter means 240-1, 240-2, 240-3 and 240-4, respectively, and further, updates the impulse responses by repeating this calculation per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

[0111]

(In case of Adaptive Type Operation)

The signals  $e_M$  and  $e_S$  outputted from the subtracters 248 and 250, assuming that frequency-axis expressions of  $e_M$  and  $e_S$  are respectively given as  $E_M$  and  $E_S$  and the filter characteristics set to the filter means 240-1, 240-2, 240-3 and 240-4 are given as  $H_{LM}$ ,  $H_{RM}$ ,  $H_{LS}$  and  $H_{RS}$  ( $h_{LM}$ ,  $h_{RM}$ ,  $h_{LS}$  and  $h_{RS}$  when expressed in terms of the impulse responses), become

$$\begin{aligned} E_M &= \{L(H_{LL}+H_{LR}) + X_R(H_{RL}+H_{RR})\} - (X_L \cdot H_{LM} + X_R \cdot H_{RM}) \\ &= L\{(H_{LL}+H_{LR}) - H_{LM}\} + X_R\{(H_{RL}+H_{RR}) - H_{RM}\} \dots (59) \end{aligned}$$

$$\begin{aligned} E_S &= \{L(H_{LL}-H_{LR}) + X_R(H_{RL}-H_{RR})\} - (X_L \cdot H_{LS} + X_R \cdot H_{RS}) \\ &= L\{(H_{LL}-H_{LR}) - H_{LS}\} + X_R\{(H_{RL}-H_{RR}) - H_{RS}\} \dots (60) \end{aligned}$$

When the composite transfer functions are given as

$$H_{LM} = H_{LL} + H_{LR}$$

$$H_{RM} = H_{RL} + H_{RR}$$

$$H_{LS} = H_{LL} - H_{LR}$$

$$H_{RS} = H_{RL} - H_{RR}$$

then

$$X_L = (X_M + X_S) / 2$$

$$X_R = (X_M - X_S) / 2$$

hence, the equations (59) and (60) respectively become

$$E_M = (X_M + X_S) \cdot (H_{LM} - H_{LM}^{\wedge}) / 2 + (X_M - X_S) \cdot (H_{RM} - H_{RM}^{\wedge}) / 2 \dots (61)$$

$$E_S = (X_M + X_S) \cdot (H_{LS} - H_{LS}^{\wedge}) / 2 + (X_M - X_S) \cdot (H_{RS} - H_{RS}^{\wedge}) / 2 \dots (62)$$

5 When the estimated errors of the composite transfer functions are given as

$$\Delta H_{LM} = H_{LM} - H_{LM}^{\wedge}$$

$$\Delta H_{RM} = H_{RM} - H_{RM}^{\wedge}$$

$$\Delta H_{LS} = H_{LS} - H_{LS}^{\wedge}$$

10 
$$\Delta H_{RS} = H_{RS} - H_{RS}^{\wedge}$$

the equations (61) and (62) respectively become

$$E_M = X_M (\Delta H_{LM} + \Delta H_{RM}) / 2 + X_S (\Delta H_{LM} - \Delta H_{RM}) / 2$$

$$E_S = X_M (\Delta H_{LS} + \Delta H_{RS}) / 2 + X_S (\Delta H_{LS} - \Delta H_{RS}) / 2$$

hence

15 
$$2E_M = X_M (\Delta H_{LM} + \Delta H_{RM}) + X_S (\Delta H_{LM} - \Delta H_{RM}) \dots (61')$$

$$2E_S = X_M (\Delta H_{LS} + \Delta H_{RS}) + X_S (\Delta H_{LS} - \Delta H_{RS}) \dots (62')$$

When both sides of the equations (61') and (62') are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$  and ensemble-averaged,

20 
$$\Sigma X_M^* \cdot 2E_M = \Sigma X_M^* \cdot X_M (\Delta H_{LM} + \Delta H_{RM}) + \Sigma X_M^* \cdot X_S (\Delta H_{LM} - \Delta H_{RM}) \dots (63)$$

$$\Sigma X_S^* \cdot 2E_M = \Sigma X_S^* \cdot X_M (\Delta H_{LM} + \Delta H_{RM}) + \Sigma X_S^* \cdot X_S (\Delta H_{LM} - \Delta H_{RM}) \dots (64)$$

25 
$$\Sigma X_M^* \cdot 2E_S = \Sigma X_M^* \cdot X_M (\Delta H_{LS} + \Delta H_{RS}) + \Sigma X_M^* \cdot X_S (\Delta H_{LS} - \Delta H_{RS}) \dots (65)$$

$$\Sigma X_S^* \cdot 2E_S = \Sigma X_S^* \cdot X_M (\Delta H_{LS} + \Delta H_{RS}) + \Sigma X_S^* \cdot X_S (\Delta H_{LS} - \Delta H_{RS}) \dots (66)$$

are respectively obtained.

[0112]

In the equations (63) to (66), since  $X_M$  and  $X_S$  are approximately uncorrelated with each other, such a term having  $X_M^* \cdot X_S$  or  $X_S^* \cdot X_M$  becomes approximately zero when ensemble-averaged. Further,

$$X_M^* \cdot X_M = |X_M|^2$$

$$X_S^* \cdot X_S = |X_S|^2$$

hence, the equations (63) to (66) respectively become

$$10 \quad \Sigma X_M^* \cdot 2E_M = \Sigma |X_M|^2 (\Delta H_{LM} + \Delta H_{RM}) \dots (63')$$

$$\Sigma X_S^* \cdot 2E_M = \Sigma |X_S|^2 (\Delta H_{LM} - \Delta H_{RM}) \dots (64')$$

$$\Sigma X_M^* \cdot 2E_S = \Sigma |X_M|^2 (\Delta H_{LS} + \Delta H_{RS}) \dots (65')$$

$$\Sigma X_S^* \cdot 2E_S = \Sigma |X_S|^2 (\Delta H_{LS} - \Delta H_{RS}) \dots (66') .$$

15 By transforming the equations (63') to (66'), the following composite transfer functions are respectively derived.

$$\Delta H_{LM} + \Delta H_{RM} = \Sigma X_M^* \cdot 2E_M / \Sigma |X_M|^2 \dots (63'')$$

$$\Delta H_{LM} - \Delta H_{RM} = \Sigma X_S^* \cdot 2E_M / \Sigma |X_S|^2 \dots (64'')$$

$$\Delta H_{LS} + \Delta H_{RS} = \Sigma X_M^* \cdot 2E_S / \Sigma |X_M|^2 \dots (65'')$$

$$20 \quad \Delta H_{LS} - \Delta H_{RS} = \Sigma X_S^* \cdot 2E_S / \Sigma |X_S|^2 \dots (66'')$$

From the equations (63'') to (66''),

$$\Delta H_{LM} = \Sigma X_M^* \cdot 2E_M / \Sigma |X_M|^2 + \Sigma X_S^* \cdot 2E_M / \Sigma |X_S|^2 \dots (67)$$

$$\Delta H_{RM} = \Sigma X_M^* \cdot 2E_M / \Sigma |X_M|^2 - \Sigma X_S^* \cdot 2E_M / \Sigma |X_S|^2 \dots (68)$$

$$\Delta H_{LS} = \Sigma X_M^* \cdot 2E_S / \Sigma |X_M|^2 + \Sigma X_S^* \cdot 2E_S / \Sigma |X_S|^2 \dots (69)$$

$$25 \quad \Delta H_{RS} = \Sigma X_M^* \cdot 2E_S / \Sigma |X_M|^2 - \Sigma X_S^* \cdot 2E_S / \Sigma |X_S|^2 \dots (70)$$

are respectively derived.

[0113]

Using the estimated errors  $\Delta H_{LM}$ ,  $\Delta H_{RM}$ ,  $\Delta H_{LS}$  and  $\Delta H_{RS}$  derived from the equations (67) to (70), the filter characteristics of the filter means 240-1, 240-2, 240-3 and 240-4 are updated per suitably determined prescribed time period (e.g. time period of performing ensemble averaging). For example, assuming that impulse responses  $h_{LM}$ ,  $h_{RM}$ ,  $h_{LS}$  and  $h_{RS}$  after K-th updating are given as  $h_{LM}(k)$ ,  $h_{RM}(k)$ ,  $h_{LS}(k)$  and  $h_{RS}(k)$ , using impulse responses  $\Delta h_{LM}$ ,  $\Delta h_{RM}$ ,  $\Delta h_{LS}$  and  $\Delta h_{RS}$  corresponding to the derived estimated errors  $\Delta H_{LM}$ ,  $\Delta H_{RM}$ ,  $\Delta H_{LS}$  and  $\Delta H_{RS}$ ,

$$h_{LM}(k+1) = h_{LM}(k) + \alpha \Delta h_{LM} \dots (71)$$

$$h_{RM}(k+1) = h_{RM}(k) + \alpha \Delta h_{RM} \dots (72)$$

$$h_{LS}(k+1) = h_{LS}(k) + \alpha \Delta h_{LS} \dots (73)$$

$$h_{RS}(k+1) = h_{RS}(k) + \alpha \Delta h_{RS} \dots (74).$$

Using these updating equations, (k+1)th impulse responses  $h_{LM}(k+1)$ ,  $h_{RM}(k+1)$ ,  $h_{LS}(k+1)$  and  $h_{RS}(k+1)$  are derived and set to the filter means 240-1, 240-2, 240-3 and 240-4, respectively, which is repeated per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

[0114]

Fig. 17 shows another structural example in the stereo echo canceller 16, 24 of Fig. 2. The same symbols are used with respect to those portions common to the foregoing structure of Fig. 16. Left/right two-channel stereo signals  $x_L$  and  $x_R$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are

outputted from sound output ends SO(L) and SO(R) as they are  
(i.e. not through sum/difference signal producing means 252),  
and reproduced at loudspeakers SP(L) and SP(R), respectively.  
The sum/difference signal producing means 252 performs  
5 addition of such stereo signals  $x_L$  and  $x_R$  using an adder 254  
so as to produce a sum signal  $x_M (= x_L + x_R)$ , while performs  
subtraction thereof using a subtracter 256 so as to produce a  
difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ .

[0115]

10 Transfer function calculating means 258 implements a  
cross-spectrum calculation between the sum signal  $x_M$  and the  
difference signal  $x_S$  produced by the sum/difference signal  
producing means 252 and signals  $e_M$  and  $e_S$  outputted from  
subtracters 248 and 250 and, based on this cross-spectrum  
15 calculation, performs setting and updating of filter  
characteristics of the filter means 240-1 to 240-4.  
Operations thereof are the same as those described with  
respect to the structure of Fig. 16. The signals  $e_M$  and  $e_S$   
outputted from the subtracters 248 and 250 are inputted into  
20 stereo audio signal demodulating means 282. The stereo audio  
signal demodulating means 282 performs addition of the  
signals  $e_M$  and  $e_S$  using an adder 284, and further, gives  
thereto a coefficient  $1/2$  using a coefficient multiplier 286  
to recover a left-channel signal  $e_L$ , while performs  
25 subtraction of the signals  $e_M$  and  $e_S$  using a subtracter 288,  
and further, gives thereto a coefficient  $1/2$  using a  
coefficient multiplier 290 to recover a right-channel signal

$e_R$ . The recovered signals  $e_L$  and  $e_R$  are respectively  
outputted from line output ends LO(L) and LO(R) and  
transmitted toward the spot on the counterpart side.  
Operations of the other portions are the same as those  
5 described with respect to the structure of Fig. 16.

[0116]

Fig. 18 shows another structural example in the  
stereo echo canceller 16, 24 of Fig. 2. The same symbols are  
used with respect to those portions common to the foregoing  
10 structure of Fig. 16 or 17. A sum signal  $x_M (= x_L + x_R)$  and a  
difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$   
transmitted from the spot on the counterpart side and  
inputted into line input ends LI(L) and LI(R) are inputted  
into stereo audio signal demodulating means 262. The stereo  
15 audio signal demodulating means 262 performs addition of the  
sum and difference signals  $x_M$  and  $x_S$  using an adder 264, and  
further, gives thereto a coefficient  $1/2$  using a coefficient  
multiplier 266 to recover the original signal  $x_L$ , while  
performs subtraction of the sum and difference signals  $x_M$  and  
20  $x_S$  using a subtracter 268, and further, gives thereto a  
coefficient  $1/2$  using a coefficient multiplier 270 to recover  
the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are  
outputted from sound output ends SO(L) and SO(R) and  
reproduced at loudspeakers SP(L) and SP(R), respectively.  
25 Sum/difference signal producing means 252 performs addition  
of such stereo signals  $x_L$  and  $x_R$  using an adder 254 so as to  
produce a sum signal  $x_M (= x_L + x_R)$ , while performs subtraction

thereof using a subtracter 256 so as to produce a difference signal  $x_s \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ .

[0117]

Transfer function calculating means 258 implements a  
5 cross-spectrum calculation between the sum signal  $x_m$  and the  
difference signal  $x_s$  produced by the sum/difference signal  
producing means 252 and signals  $e_m$  and  $e_s$  outputted from  
subtracters 248 and 250 and, based on this cross-spectrum  
calculation, performs setting and updating of filter  
10 characteristics of filter means 240-1 to 240-4. Operations  
thereof are the same as those described with respect to the  
structure of Fig. 16 or 17. The signals  $e_m$  and  $e_s$  outputted  
from the subtracters 248 and 250 are respectively outputted  
from line output ends LO(L) and LO(R) and transmitted toward  
15 the spot on the counterpart side. Operations of the other  
portions are the same as those described with respect to the  
structure of Fig. 16 or 17.

[0118]

Fig. 19 shows another structural example in the  
20 stereo echo canceller 16, 24 of Fig. 2. The same symbols are  
used with respect to those portions common to the foregoing  
structure of Fig. 16, 17 or 18. Left/right two-channel  
stereo signals  $x_L$  and  $x_R$  transmitted from the spot on the  
counterpart side and inputted into line input ends LI(L) and  
25 LI(R) are inputted into sum/difference signal producing means  
252. The sum/difference signal producing means 252 performs  
addition of the stereo signals  $x_L$  and  $x_R$  using an adder 254

so as to produce a sum signal  $x_M (= x_L + x_R)$ , while performs subtraction thereof using a subtracter 256 so as to produce a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ . The produced sum signal  $x_M$  and difference signal  $x_S$  are inputted  
5 into stereo audio signal demodulating means 262. The stereo audio signal demodulating means 262 performs addition of the sum and difference signals  $x_M$  and  $x_S$  using an adder 264, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 266 to recover the original signal  $x_L$ , while  
10 performs subtraction of the sum and difference signals  $x_M$  and  $x_S$  using a subtracter 268, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 270 to recover the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are outputted from sound output ends  $SO(L)$  and  $SO(R)$  and  
15 reproduced at loudspeakers  $SP(L)$  and  $SP(R)$ , respectively.

[0119]

Transfer function calculating means 258 implements a cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal  
20 producing means 252 and signals  $e_M$  and  $e_S$  outputted from subtracters 248 and 250 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics of the filter means 240-1 to 240-4. Operations thereof are the same as those described with  
25 respect to the structure of Fig. 16, 17 or 18. The signals  $e_M$  and  $e_S$  outputted from the subtracters 248 and 250 are inputted into stereo audio signal demodulating means 282.

The stereo audio signal demodulating means 282 performs addition of the signals  $e_M$  and  $e_S$  using an adder 284, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 286 to recover a left-channel signal  $e_L$ , while  
5 performs subtraction of the signals  $e_M$  and  $e_S$  using a subtracter 288, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 290 to recover a right-channel signal  $e_R$ . The recovered signals  $e_L$  and  $e_R$  are respectively outputted from line output ends LO(L) and LO(R) and  
10 transmitted toward the spot on the counterpart side. Operations of the other portions are the same as those described with respect to the structure of Fig. 16, 17 or 18.

[0120]

Fig. 20 shows another structural example in the  
15 stereo echo canceller 16, 24. A sum signal  $x_M (= x_L + x_R)$  and a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are inputted into stereo audio signal demodulating means 362. The stereo  
20 audio signal demodulating means 362 performs addition of the sum and difference signals  $x_M$  and  $x_S$  using an adder 364, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 366 to recover the original signal  $x_L$ , while performs subtraction of the sum and difference signals  $x_M$  and  
25  $x_S$  using a subtracter 368, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 370 to recover the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are

outputted from sound output ends SO(L) and SO(R) and reproduced at loudspeakers SP(L) and SP(R), respectively.

[0121]

Collected audio signals  $y_L$  and  $y_R$  of microphones MC(L) and MC(R) are inputted into sum/difference signal producing means 372. The sum/difference signal producing means 372 implements addition of the microphone collected audio signals  $y_L$  and  $y_R$  using an adder 373 so as to produce a sum signal  $y_M$ , while implements subtraction thereof using a subtracter 375 so as to produce a difference signal  $y_s$ .

[0122]

Filer means 340-1 to 340-4 are formed by, for example, FIR filters. These filter means 340-1 to 340-4 are each set with an impulse response corresponding to a composite transfer function in the form of combination of transfer functions  $H_{LL}$ ,  $H_{LR}$ ,  $H_{RL}$  and  $H_{RR}$  of four audio transfer systems between the loudspeakers SP(L) and SP(R) and microphones MC(L) and MC(R), respectively, and perform, using such impulse responses, a convolution calculation of the sum signal  $x_M$  and the difference signal  $x_s$  inputted from the line input ends LI(L) and LI(R), thereby producing echo cancel signals EC1 to EC4, respectively.

[0123]

An adder 344 performs a calculation of  $EC1+EC3$ . An adder 346 performs a calculation of  $EC2+EC4$ . A subtracter 348 subtracts an echo cancel signal  $EC1+EC3$  from the sum signal  $y_M$ , thereby to perform echo cancellation. A

subtractor 350 subtracts an echo cancel signal  $EC2+EC4$  from the difference signal  $y_s$ , thereby to perform echo cancellation. Signals  $e_m$  and  $e_s$  outputted from the subtracters 348 and 350 are outputted from line output ends LO(L) and LO(R), respectively, and transmitted toward the spot on the counterpart side.

[0124]

Transfer function calculating means 358 implements a cross-spectrum calculation between the sum signal  $x_m$  and the difference signal  $x_s$  inputted from the line input ends LI(L) and LI(R) and the signals  $e_m$  and  $e_s$  outputted from the subtracters 348 and 350 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics (impulse responses) of the filter means 340-1 to 340-4. Specifically, upon starting the system, the filter characteristics of the filter means 340-1 to 340-4 are not set, i.e. coefficients are all set to zero, so that the echo cancel signals  $EC1$  to  $EC4$  are zero, and thus the sum signal  $y_m$  and the difference signal  $y_s$  outputted from the sum/difference signal producing means 372, as they are, are outputted from the subtracters 348 and 350. Therefore, at this time, the transfer function calculating means 358 performs the cross-spectrum calculation between the sum signal  $x_m$  and the difference signal  $x_s$  inputted from the line input ends LI(L) and LI(R) and the sum signal  $e_m$  and the difference signal  $e_s$  outputted from the subtracters 348 and 350 and, based on this cross-spectrum calculation, derives a

plurality of composite transfer functions each in the form of combination of transfer functions  $H_{LL}$ ,  $H_{LR}$ ,  $H_{RL}$  and  $H_{RR}$  of four audio transfer systems between the loudspeakers SP(L) and SP(R) and the microphones MC(L) and MC(R), respectively, and  
5 implements initial setting of the filter characteristics of the filter means 340-1 to 340-4 to values corresponding to such composite transfer functions. After the initial setting, since the echo cancel signals are produced by the filter means 340-1 to 340-4, the echo cancel error signals  $e_M$  and  $e_S$   
10 corresponding to difference signals between the sum signal  $y_M$  and the difference signal  $y_S$  outputted from the sum/difference signal producing means 372 and the echo cancel signals EC1 to EC4 are outputted from the subtracters 348 and 350. Therefore, at this time, the transfer function  
15 calculating means 358 performs the cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  inputted from the line input ends LI(L) and LI(R) and the echo cancel error signals  $e_M$  and  $e_S$  outputted from the subtracters 348 and 350 and, based on this cross-spectrum  
20 calculation, derives estimated errors of the foregoing composite transfer functions, respectively, and updates the filter characteristics of the filter means 340-1 to 340-4 to values that cancel such estimated errors, respectively. By repeating this updating operation per prescribed time period,  
25 the echo cancel error can be converged to a minimum value. Further, even if the transfer functions change due to movement of the microphone positions or the like, the echo

cancel error can be converged to a minimum value by sequentially updating the filter characteristics of the filter means 340-1 to 340-4 depending thereon.

[0125]

5           Correlation detecting means 360 detects a correlation between the sum signal  $x_M$  and the difference signal  $x_S$  based on a correlation value calculation or the like, and stops updating of the foregoing filter characteristics when the correlation value is no less than a  
10           prescribed value. When the correlation value becomes lower than the prescribed value, updating of the foregoing filter characteristics is restarted.

[0126]

          Herein, the filter characteristics (impulse  
15           responses) that are set to the filter means 340-1 to 340-4 by the transfer function calculating means 358 will be described. In the transfer function calculating means 358, the following calculation is performed.

(In case of Fixed Type Operation)

20           The sum and difference signals  $y_M$  and  $y_S$  of the output signals  $y_L$  and  $y_R$  of the microphones MC(L) and MC(R) become

$$\begin{aligned} Y_M &= Y_L + Y_R \\ &= (X_L \cdot H_{LL} + X_R \cdot H_{RL}) + (X_L \cdot H_{LR} + X_R \cdot H_{RR}) \\ 25 \quad &= X_L (H_{LL} + H_{LR}) + X_R (H_{RL} + H_{RR}) \quad \dots (71) \\ Y_S &= Y_L - Y_R \\ &= (X_L \cdot H_{LL} + X_R \cdot H_{RL}) - (X_L \cdot H_{LR} + X_R \cdot H_{RR}) \end{aligned}$$

$$= X_L (H_{LL} - H_{LR}) + X_R (H_{RL} - H_{RR}) \dots (72).$$

$$X_L = (X_M + X_S) / 2$$

$$X_R = (X_M - X_S) / 2$$

hence, the equations (71) and (72) respectively become

$$\begin{aligned} 5 \quad Y_M &= (X_M + X_S) \cdot (H_{LL} + H_{LR}) / 2 + (X_M - X_S) \cdot (H_{RL} + H_{RR}) / 2 \\ &= X_M (H_{LL} + H_{LR} + H_{RL} + H_{RR}) / 2 + X_S (H_{LL} + H_{LR} - H_{RL} - H_{RR}) / 2 \dots (71') \\ Y_S &= (X_M + X_S) \cdot (H_{LL} - H_{LR}) / 2 + (X_M - X_S) \cdot (H_{RL} - H_{RR}) / 2 \\ &= X_M (H_{LL} - H_{LR} + H_{RL} - H_{RR}) / 2 + X_S (H_{LL} - H_{LR} - H_{RL} + H_{RR}) / 2 \dots (72') \end{aligned}$$

When the composite transfer functions are given as

$$\begin{aligned} 10 \quad H_{MM} &= (H_{LL} + H_{LR} + H_{RL} + H_{RR}) / 2 \\ H_{SM} &= (H_{LL} + H_{LR} - H_{RL} - H_{RR}) / 2 \\ H_{MS} &= (H_{LL} - H_{LR} + H_{RL} - H_{RR}) / 2 \\ H_{SS} &= (H_{LL} - H_{LR} - H_{RL} + H_{RR}) / 2 \end{aligned}$$

the equations (71') and (72') respectively become

$$\begin{aligned} 15 \quad Y_M &= X_M \cdot H_{MM} + X_S \cdot H_{SM} \dots (71'') \\ Y_S &= X_M \cdot H_{MS} + X_S \cdot H_{SS} \dots (72''). \end{aligned}$$

When both sides of the equations (71'') and (72'') are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$  and ensemble-averaged,

$$20 \quad \Sigma X_M^* \cdot Y_M = \Sigma X_M^* \cdot X_M \cdot H_{MM} + \Sigma X_M^* \cdot X_S \cdot H_{SM} \dots (73)$$

$$\Sigma X_S^* \cdot Y_M = \Sigma X_S^* \cdot X_M \cdot H_{MM} + \Sigma X_S^* \cdot X_S \cdot H_{SM} \dots (74)$$

$$\Sigma X_M^* \cdot Y_S = \Sigma X_M^* \cdot X_M \cdot H_{MS} + \Sigma X_M^* \cdot X_S \cdot H_{SS} \dots (75)$$

$$\Sigma X_S^* \cdot Y_S = \Sigma X_S^* \cdot X_M \cdot H_{MS} + \Sigma X_S^* \cdot X_S \cdot H_{SS} \dots (76)$$

are respectively obtained.

$$25 \quad [0127]$$

In the equations (73) to (76), since  $X_M$  and  $X_S$  are approximately uncorrelated with each other, such a term

having  $X_M^* \cdot X_S$  or  $X_S^* \cdot X_M$  becomes approximately zero when ensemble-averaged. Further,

$$X_M^* \cdot X_M = |X_M|^2$$

$$X_S^* \cdot X_S = |X_S|^2$$

5 hence, the equations (73) to (76) respectively become

$$\Sigma X_M^* \cdot Y_M = \Sigma |X_M|^2 \cdot H_{MM} \dots (73')$$

$$\Sigma X_S^* \cdot Y_M = \Sigma |X_S|^2 \cdot H_{SM} \dots (74')$$

$$\Sigma X_M^* \cdot Y_S = \Sigma |X_M|^2 \cdot H_{MS} \dots (75')$$

$$\Sigma X_S^* \cdot Y_S = \Sigma |X_S|^2 \cdot H_{SS} \dots (76').$$

10 From the equations (73') to (76'),

$$H_{MM} = \Sigma X_M^* \cdot Y_M / \Sigma |X_M|^2 \dots (77)$$

$$H_{SM} = \Sigma X_S^* \cdot Y_M / \Sigma |X_S|^2 \dots (78)$$

$$H_{MS} = \Sigma X_M^* \cdot Y_S / \Sigma |X_M|^2 \dots (79)$$

$$H_{SS} = \Sigma X_S^* \cdot Y_S / \Sigma |X_S|^2 \dots (80)$$

15 are respectively derived.

[0128]

Impulse responses  $h_{MM}$ ,  $h_{SM}$ ,  $h_{MS}$  and  $h_{SS}$  obtained by applying the inverse Fourier transformation to these derived composite transfer functions  $H_{MM}$ ,  $H_{SM}$ ,  $H_{MS}$  and  $H_{SS}$  are the  
20 filter characteristics to be set to the filter means 340-1, 340-2, 340-3 and 340-4, respectively. Therefore, the transfer function calculating means 358 derives the respective composite transfer functions  $H_{MM}$ ,  $H_{SM}$ ,  $H_{MS}$  and  $H_{SS}$  from the equations (77) to (80) based on the sum signal  $x_M$   
25 and the difference signal  $x_S$  inputted into the line input ends LI(L) and LI(R) and the sum signal  $y_M$  and the difference signal  $y_S$  outputted from the sum/difference signal producing

means 372, derives the impulse responses  $h_{MM}$ ,  $h_{SM}$ ,  $h_{MS}$  and  $h_{SS}$  by applying the inverse Fourier transformation to those derived composite transfer functions, sets the derived impulse responses to the filter means 340-1, 340-2, 340-3 and 340-4, respectively, and further, updates the impulse responses by repeating this calculation per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

[0129]

10 (In case of Adaptive Type Operation)

Assuming that the filter characteristics set to the filter means 340-1, 340-2, 340-3 and 340-4 are given as  $H^{MM}$ ,  $H^{SM}$ ,  $H^{MS}$  and  $H^{SS}$  ( $h^{MM}$ ,  $h^{SM}$ ,  $h^{MS}$  and  $h^{SS}$  when expressed in terms of the impulse responses), the signals  $e_M$  and  $e_S$  outputted from the subtracters 348 and 350 become

$$\begin{aligned} E_M &= \{ (X_M + X_S) \cdot (H_{LL} + H_{LR}) / 2 + (X_M - X_S) \cdot (H_{RL} + H_{RR}) / 2 \} \\ &\quad - (X_M \cdot H^{MM} + X_S \cdot H^{SM}) \\ &= X_M [ \{ (H_{LL} + H_{LR} + H_{RL} + H_{RR}) / 2 \} - H^{MM} ] \\ &\quad + X_S [ \{ (H_{LL} + H_{LR} - H_{RL} - H_{RR}) / 2 \} - H^{SM} ] \dots (81) \end{aligned}$$

$$\begin{aligned} E_S &= \{ (X_M + X_S) \cdot (H_{LL} - H_{LR}) / 2 + (X_M - X_S) \cdot (H_{RL} - H_{RR}) / 2 \} \\ &\quad - (X_M \cdot H^{MS} + X_S \cdot H^{SS}) \\ &= X_M [ \{ (H_{LL} - H_{LR} + H_{RL} - H_{RR}) / 2 \} - H^{MS} ] \\ &\quad + X_S [ \{ (H_{LL} - H_{LR} - H_{RL} + H_{RR}) / 2 \} - H^{SS} ] \dots (82) \end{aligned}$$

When the composite transfer functions are given as

$$H_{MM} = (H_{LL} + H_{LR} + H_{RL} + H_{RR}) / 2$$

$$H_{SM} = (H_{LL} + H_{LR} - H_{RL} - H_{RR}) / 2$$

$$H_{MS} = (H_{LL} - H_{LR} + H_{RL} - H_{RR}) / 2$$

$$H_{SS} = (H_{LL} - H_{LR} - H_{RL} + H_{RR}) / 2$$

the equations (81) and (82) respectively become

$$E_M = X_M (H_{MM} - \hat{H}_{MM}) + X_S (H_{SM} - \hat{H}_{SM}) \dots (81')$$

$$E_S = X_M (H_{MS} - \hat{H}_{MS}) + X_S (H_{SS} - \hat{H}_{SS}) \dots (82')$$

5

When

$$\Delta H_{MM} = H_{MM} - \hat{H}_{MM}$$

$$\Delta H_{SM} = H_{SM} - \hat{H}_{SM}$$

$$\Delta H_{MS} = H_{MS} - \hat{H}_{MS}$$

$$\Delta H_{SS} = H_{SS} - \hat{H}_{SS}$$

10

are given, the equations (81') and (82') respectively become

$$E_M = X_M \cdot \Delta H_{MM} + X_S \cdot \Delta H_{SM} \dots (81'')$$

$$E_S = X_M \cdot \Delta H_{MS} + X_S \cdot \Delta H_{SS} \dots (82'')$$

When both sides of the equations (81'') and (82'') are multiplied by complex conjugates  $X_M^*$  and  $X_S^*$  of  $X_M$  and  $X_S$  and ensemble-averaged,

15

$$\Sigma X_M^* \cdot E_M = \Sigma X_M^* \cdot X_M \cdot \Delta H_{MM} + \Sigma X_M^* \cdot X_S \cdot \Delta H_{SM} \dots (83)$$

$$\Sigma X_S^* \cdot E_M = \Sigma X_S^* \cdot X_M \cdot \Delta H_{MM} + \Sigma X_S^* \cdot X_S \cdot \Delta H_{SM} \dots (84)$$

$$\Sigma X_M^* \cdot E_S = \Sigma X_M^* \cdot X_M \cdot \Delta H_{MS} + \Sigma X_M^* \cdot X_S \cdot \Delta H_{SS} \dots (85)$$

$$\Sigma X_S^* \cdot E_S = \Sigma X_S^* \cdot X_M \cdot \Delta H_{MS} + \Sigma X_S^* \cdot X_S \cdot \Delta H_{SS} \dots (86)$$

20

are respectively obtained.

[0130]

In the equations (83) to (86), since  $X_M$  and  $X_S$  are approximately uncorrelated with each other, such a term having  $X_M^* \cdot X_S$  or  $X_S^* \cdot X_M$  becomes approximately zero when ensemble-averaged. Further,

25

$$X_M^* \cdot X_M = |X_M|^2$$

$$X_S^* \cdot X_S = |X_S|^2$$

hence, the equations (83) to (86) respectively become

$$\Sigma X_M^* \cdot E_M = \Sigma |X_M|^2 \cdot \Delta H_{MM} \dots (83')$$

$$\Sigma X_S^* \cdot E_M = \Sigma |X_S|^2 \cdot \Delta H_{SM} \dots (84')$$

$$\Sigma X_M^* \cdot E_S = \Sigma |X_M|^2 \cdot \Delta H_{MS} \dots (85')$$

5  $\Sigma X_S^* \cdot E_S = \Sigma |X_S|^2 \cdot \Delta H_{SS} \dots (86') .$

From the equations (83') to (86'),

$$\Delta H_{MM} = \Sigma X_M^* \cdot E_M / \Sigma |X_M|^2 \dots (87)$$

$$\Delta H_{SM} = \Sigma X_S^* \cdot E_M / \Sigma |X_S|^2 \dots (88)$$

$$\Delta H_{MS} = \Sigma X_M^* \cdot E_S / \Sigma |X_M|^2 \dots (89)$$

10  $\Delta H_{SS} = \Sigma X_S^* \cdot E_S / \Sigma |X_S|^2 \dots (90)$

are respectively derived.

[0131]

Using the estimated errors  $\Delta H_{MM}$ ,  $\Delta H_{SM}$ ,  $\Delta H_{MS}$  and  $\Delta H_{SS}$  derived from the equations (87) to (90), the filter

15 characteristics of the filter means 340-1, 340-2, 340-3 and 340-4 are updated per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

For example, assuming that impulse responses  $h_{MM}$ ,  $h_{SM}$ ,  $h_{MS}$  and  $h_{SS}$  after K-th updating are given as  $h_{MM}(k)$ ,  $h_{SM}(k)$ ,  $h_{MS}(k)$  and

20  $h_{SS}(k)$ , using impulse responses  $\Delta h_{MM}$ ,  $\Delta h_{SM}$ ,  $\Delta h_{MS}$  and  $\Delta h_{SS}$

corresponding to the derived estimated errors  $\Delta H_{MM}$ ,  $\Delta H_{SM}$ ,  $\Delta H_{MS}$  and  $\Delta H_{SS}$ ,

$$h_{MM}(k+1) = h_{MM}(k) + \alpha \Delta h_{MM} \dots (91)$$

$$h_{SM}(k+1) = h_{SM}(k) + \alpha \Delta h_{SM} \dots (92)$$

25  $h_{MS}(k+1) = h_{MS}(k) + \alpha \Delta h_{MS} \dots (93)$

$$h_{SS}(k+1) = h_{SS}(k) + \alpha \Delta h_{SS} \dots (94) .$$

Using these updating equations, (k+1)th impulse

responses  $h_{MM}(k+1)$ ,  $h_{SM}(k+1)$ ,  $h_{MS}(k+1)$  and  $h_{SS}(k+1)$  are derived and set to the filter means 340-1, 340-2, 340-3 and 340-4, respectively, which is repeated per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

[0132]

Fig. 21 shows another structural example in the stereo echo canceller 16, 24 of Fig. 2. The same symbols are used with respect to those portions common to the foregoing structure of Fig. 20. Left/right two-channel stereo signals  $x_L$  and  $x_R$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are outputted from sound output ends SO(L) and SO(R) as they are (i.e. not through sum/difference signal producing means 352), and reproduced at loudspeakers SP(L) and SP(R), respectively. The sum/difference signal producing means 352 performs addition of such stereo signals  $x_L$  and  $x_R$  using an adder 354 so as to produce a sum signal  $x_M (= x_L + x_R)$ , while performs subtraction thereof using a subtracter 356 so as to produce a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ .

[0133]

Transfer function calculating means 358 implements a cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal producing means 352 and signals  $e_M$  and  $e_S$  outputted from subtracters 348 and 350 and, based on this cross-spectrum calculation, performs setting and updating of filter

characteristics of the filter means 340-1 to 340-4.

Operations thereof are the same as those described with

respect to the structure of Fig. 20. The signals  $e_M$  and  $e_S$

outputted from the subtracters 348 and 350 are inputted into

5 stereo audio signal demodulating means 382. The stereo audio

signal demodulating means 382 performs addition of the

signals  $e_M$  and  $e_S$  using an adder 384, and further, gives

thereto a coefficient  $1/2$  using a coefficient multiplier 386

to recover a left-channel signal  $e_L$ , while performs

10 subtraction of the signals  $e_M$  and  $e_S$  using a subtracter 388,

and further, gives thereto a coefficient  $1/2$  using a

coefficient multiplier 390 to recover a right-channel signal

$e_R$ . The recovered signals  $e_L$  and  $e_R$  are respectively

outputted from line output ends LO(L) and LO(R) and

15 transmitted toward the spot on the counterpart side.

Operations of the other portions are the same as those

described with respect to the structure of Fig. 20.

[0134]

Fig. 22 shows another structural example in the

20 stereo echo canceller 16, 24 of Fig. 2. The same symbols are

used with respect to those portions common to the foregoing

structure of Fig. 20 or 21. A sum signal  $x_M (= x_L + x_R)$  and a

difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L) \}$

transmitted from the spot on the counterpart side and

25 inputted into line input ends LI(L) and LI(R) are inputted

into stereo audio signal demodulating means 362. The stereo

audio signal demodulating means 362 performs addition of the

sum and difference signals  $x_M$  and  $x_S$  using an adder 364, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 366 to recover the original signal  $x_L$ , while performs subtraction of the sum and difference signals  $x_M$  and  $x_S$  using a subtracter 368, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 370 to recover the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are outputted from sound output ends  $SO(L)$  and  $SO(R)$  and reproduced at loudspeakers  $SP(L)$  and  $SP(R)$ , respectively.

Sum/difference signal producing means 352 performs addition of such stereo signals  $x_L$  and  $x_R$  using an adder 354 so as to produce a sum signal  $x_M (= x_L + x_R)$ , while performs subtraction thereof using a subtracter 356 so as to produce a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ .

[0135]

Transfer function calculating means 358 implements a cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal producing means 352 and signals  $e_M$  and  $e_S$  outputted from subtracters 348 and 350 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics of filter means 340-1 to 340-4. Operations thereof are the same as those described with respect to the structure of Fig. 20 or 21. The signals  $e_M$  and  $e_S$  outputted from the subtracters 348 and 350 are respectively outputted from line output ends  $LO(L)$  and  $LO(R)$  and transmitted toward the spot on the counterpart side. Operations of the other

portions are the same as those described with respect to the structure of Fig. 20 or 21.

[0136]

Fig. 23 shows another structural example in the stereo echo canceller 16, 24 of Fig. 2. The same symbols are used with respect to those portions common to the foregoing structure of Fig. 20, 21 or 22. Left/right two-channel stereo signals  $x_L$  and  $x_R$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are inputted into sum/difference signal producing means 352. The sum/difference signal producing means 352 performs addition of the stereo signals  $x_L$  and  $x_R$  using an adder 354 so as to produce a sum signal  $x_M (= x_L + x_R)$ , while performs subtraction thereof using a subtracter 356 so as to produce a difference signal  $x_S \{= x_L - x_R \text{ (or it may also be } x_R - x_L)\}$ . The produced sum signal  $x_M$  and difference signal  $x_S$  are inputted into stereo audio signal demodulating means 362. The stereo audio signal demodulating means 362 performs addition of the sum and difference signals  $x_M$  and  $x_S$  using an adder 364, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 366 to recover the original signal  $x_L$ , while performs subtraction of the sum and difference signals  $x_M$  and  $x_S$  using a subtracter 368, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 370 to recover the original signal  $x_R$ . The recovered signals  $x_L$  and  $x_R$  are outputted from sound output ends SO(L) and SO(R) and reproduced at loudspeakers SP(L) and SP(R), respectively.

[0137]

Transfer function calculating means 358 implements a cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal producing means 352 and signals  $e_M$  and  $e_S$  outputted from subtracters 348 and 350 and, based on this cross-spectrum calculation, performs setting and updating of filter characteristics of the filter means 340-1 to 340-4.

Operations thereof are the same as those described with respect to the structure of Fig. 20, 21 or 22. The signals  $e_M$  and  $e_S$  outputted from the subtracters 348 and 350 are inputted into stereo audio signal demodulating means 382. The stereo audio signal demodulating means 382 performs addition of the signals  $e_M$  and  $e_S$  using an adder 384, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 386 to recover a left-channel signal  $e_L$ , while performs subtraction of the signals  $e_M$  and  $e_S$  using a subtracter 388, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 390 to recover a right-channel signal  $e_R$ . The recovered signals  $e_L$  and  $e_R$  are respectively outputted from line output ends LO(L) and LO(R) and transmitted toward the spot on the counterpart side. Operations of the other portions are the same as those described with respect to the structure of Fig. 20, 21 or 22.

[0138]

This invention can take various structures in addition to the structures as shown in the foregoing

embodiments. For example, part of the structures of Figs. 1, 9 to 11 can be changed like Fig. 24. Specifically, sum/difference signal producing means 402 is disposed between the adders 44 and 46 and the subtracters 48 and 50, so as to perform addition of the echo cancel signals  $EC1+EC3$  and  $EC2+EC4$  using an adder 404 to produce an echo cancel signal  $(EC1+EC3)+(EC2+EC4)$ , while perform subtraction thereof using a subtracter 406 to produce  $(EC1+EC3)-(EC2+EC4)$ . Further, sum/difference signal producing means 408 is disposed between the sound input ends  $SI(L)$  and  $SI(R)$  and the subtracters 48 and 50, so as to perform addition of the collected audio signals  $y_L$  and  $y_R$  of the microphones  $MC(L)$  and  $MC(R)$  using an adder 410 to produce a sum signal  $y_M$ , while perform subtraction thereof using a subtracter 412 to produce a difference signal  $y_S$ . The subtracter 48 subtracts the echo cancel signal  $(EC1+EC3)+(EC2+EC4)$  from the sum signal  $y_M$  to implement echo cancellation. The subtracter 50 subtracts the echo cancel signal  $(EC1+EC3)-(EC2+EC4)$  from the difference signal  $y_S$  to implement echo cancellation. The signals  $e_M$  and  $e_S$  outputted from the subtracters 48 and 50 are inputted into stereo audio signal demodulating means 414. The stereo audio signal demodulating means 414 performs addition of the signals  $e_M$  and  $e_S$  using an adder 416, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 418 to recover a left-channel signal  $e_L$ , while performs subtraction of the signals  $e_M$  and  $e_S$  using a subtracter 420, and further, gives thereto a coefficient  $1/2$  using a

coefficient multiplier 422 to recover a right-channel signal  $e_R$ . The other portions are the same as those described with respect to the structures of Figs. 1, 9 to 11.

[0139]

5           On the other hand, part of the structures of Figs. 12 to 15 can be changed like Fig. 25. Specifically, sum/difference signal producing means 432 is disposed between the adders 144 and 146 and the subtracters 148 and 150, so as to perform addition of the echo cancel signals  $EC1+EC3$  and  
10  $EC2+EC4$  using an adder 434 to produce an echo cancel signal  $(EC1+EC3)+(EC2+EC4)$ , while perform subtraction thereof using a subtracter 436 to produce  $(EC1+EC3)-(EC2+EC4)$ . Further, sum/difference signal producing means 438 is disposed between the sound input ends  $SI(L)$  and  $SI(R)$  and the subtracters 148  
15 and 150, so as to perform addition of the collected audio signals  $y_L$  and  $y_R$  of the microphones  $MC(L)$  and  $MC(R)$  using an adder 440 to produce a sum signal  $y_M$ , while perform subtraction thereof using a subtracter 442 to produce a difference signal  $y_S$ . The subtracter 148 subtracts the echo  
20 cancel signal  $(EC1+EC3)+(EC2+EC4)$  from the sum signal  $y_M$  to implement echo cancellation. The subtracter 150 subtracts the echo cancel signal  $(EC1+EC3)-(EC2+EC4)$  from the difference signal  $y_S$  to implement echo cancellation. The signals  $e_M$  and  $e_S$  outputted from the subtracters 148 and 150  
25 are inputted into stereo audio signal demodulating means 444. The stereo audio signal demodulating means 444 performs addition of the signals  $e_M$  and  $e_S$  using an adder 446, and

further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 448 to recover a left-channel signal  $e_L$ , while performs subtraction of the signals  $e_M$  and  $e_S$  using a subtracter 450, and further, gives thereto a coefficient  $1/2$  using a coefficient multiplier 452 to recover a right-channel signal  $e_R$ . The other portions are the same as those described with respect to the structures of Figs. 12 to 15.

[0140]

On the other hand, part of the structures of Figs. 16 to 19 can be changed like Fig. 26. Specifically, sum/difference signal producing means 462 is disposed between the adders 244 and 246 and the subtracters 248 and 250, so as to perform addition of the echo cancel signals  $EC1+EC3$  and  $EC2+EC4$  using an adder 464 to produce an echo cancel signal  $(EC1+EC3)+(EC2+EC4)$ , while perform subtraction thereof using a subtracter 466 to produce  $(EC1+EC3)-(EC2+EC4)$ . The sum/difference signal producing means 272 of Figs. 16 to 19 is not required. The subtracter 248 subtracts the echo cancel signal  $(EC1+EC3)+(EC2+EC4)$  from the collected audio signal  $y_L$  of the microphone  $MC(L)$  to implement echo cancellation. The subtracter 250 subtracts the echo cancel signal  $(EC1+EC3)-(EC2+EC4)$  from the collected audio signal  $y_R$  of the microphone  $MC(R)$  to implement echo cancellation. Sum/difference signal producing means 468 is disposed on the output side of the subtracters 248 and 250, so as to perform addition of the output signals  $e_L$  and  $e_R$  of the subtracters 248 and 250 using an adder 470 to produce a sum signal  $e_M$ ,

while perform subtraction thereof using a subtracter 472 to produce a difference signal  $e_s$ . The other portions are the same as those described with respect to the structures of Figs. 16 to 19.

5 [0141]

On the other hand, part of the structures of Figs. 20 and 21 can be changed like Fig. 27. Specifically, sum/difference signal producing means 482 is disposed between the adders 344 and 346 and the subtracters 348 and 350, so as to perform addition of the echo cancel signals  $EC1+EC3$  and  $EC2+EC4$  using an adder 484 to produce an echo cancel signal  $(EC1+EC3)+(EC2+EC4)$ , while perform subtraction thereof using a subtracter 486 to produce  $(EC1+EC3)-(EC2+EC4)$ . The sum/difference signal producing means 372 of Figs. 20 to 23 is not required. The subtracter 348 subtracts the echo cancel signal  $(EC1+EC3)+(EC2+EC4)$  from the collected audio signal  $y_L$  of the microphone  $MC(L)$  to implement echo cancellation. The subtracter 350 subtracts the echo cancel signal  $(EC1+EC3)-(EC2+EC4)$  from the collected audio signal  $y_R$  of the microphone  $MC(R)$  to implement echo cancellation. Sum/difference signal producing means 488 is disposed on the output side of the subtracters 348 and 350, so as to perform addition of the output signals  $e_L$  and  $e_R$  of the subtracters 348 and 350 using an adder 490 to produce a sum signal  $e_M$ , while perform subtraction thereof using a subtracter 492 to produce a difference signal  $e_s$ . The other portions are the same as those described with respect to the structures of

Figs. 20 to 23.

[0142]

Fig. 28 shows another structural example in the stereo echo canceller 16, 24 of Fig. 2. The same symbols are used with respect to those portions common to the foregoing structure of Fig. 1. Left/right two-channel stereo signals  $x_L$  and  $x_R$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are outputted from sound output ends SO(L) and SO(R) as they are (i.e. not through an orthogonalizing filter 500), and reproduced at loudspeakers SP(L) and SP(R), respectively.

[0143]

Filter means 40-1 is set with an impulse response corresponding to a transfer function between the loudspeaker SP(L) and a microphone MC(L), filter means 40-2 is set with an impulse response corresponding to a transfer function between the loudspeaker SP(L) and a microphone MC(R), filter means 40-3 is set with an impulse response corresponding to a transfer function between the loudspeaker SP(R) and the microphone MC(L), and filter means 40-4 is set with an impulse response corresponding to a transfer function between the loudspeaker SP(R) and the microphone MC(R).

[0144]

The orthogonalizing filter 500 performs a principal component analysis with respect to the input stereo signals  $x_L$  and  $x_R$  per prescribed time period, and converts such input stereo signals  $x_L$  and  $x_R$  into two signals that are orthogonal

to each other. Transfer function calculating means 502 implements a cross-spectrum calculation between the mutually orthogonal two signals produced at the orthogonalizing filter 500 and signals  $e_L$  and  $e_R$  outputted from subtracters 48 and 50 and, based on this cross-spectrum calculation, sets filter characteristics (impulse responses) of the filter means 40-1 to 40-4. Specifically, upon starting the system, the filter characteristics of the filter means 40-1 to 40-4 are not set, i.e. coefficients are all set to zero, so that echo cancel signals EC1 to EC4 are zero, and thus collected audio signals of the microphones MC(L) and MC(R) themselves are outputted from the subtracters 48 and 50. Therefore, at this time, the transfer function calculating means 502 performs the cross-spectrum calculation between the mutually orthogonal two signals produced at the orthogonalizing filter 500 and the collected audio signals  $e_L$  and  $e_R$  of the microphones MC(L) and MC(R) outputted from the subtracters 48 and 50 and, based on this cross-spectrum calculation, derives transfer functions of four audio transfer systems between the loudspeakers SP(L) and SP(R) and the microphones MC(L) and MC(R), respectively, and implements initial setting of the filter characteristics of the filter means 40-1 to 40-4 to values corresponding to such transfer functions. After the initial setting, since the echo cancel signals are produced by the filter means 40-1 to 40-4, the echo cancel error signals  $e_L$  and  $e_R$  corresponding to difference signals between the collected audio signals of the microphones MC(L) and

MC(R) and the echo cancel signals EC1 to EC4 are outputted from the subtracters 48 and 50. Therefore, at this time, the transfer function calculating means 502 performs the cross-spectrum calculation between the mutually orthogonal two  
5 signals produced at the orthogonalizing filter 500 and the echo cancel error signals  $e_L$  and  $e_R$  outputted from the subtracters 48 and 50 and, based on this cross-spectrum calculation, derives estimated errors of the transfer functions of the four audio transfer systems between the  
10 loudspeakers SP(L) and SP(R) and the microphones MC(L) and MC(R), respectively, and updates the filter characteristics of the filter means 40-1 to 40-4 to values that cancel the estimated errors, respectively. By repeating this updating operation per prescribed time period, the echo cancel error  
15 can be converged to a minimum value. Further, even if the transfer functions change due to movement of the microphone positions or the like, the echo cancel error can be converged to a minimum value by sequentially updating the filter characteristics of the filter means 40-1 to 40-4 depending  
20 thereon.

[0145]

According to the known technique based on comparison between the left/right two-channel stereo signals  $x_L$  and  $x_R$  inputted into the line input ends LI(L) and LI(R) and the  
25 signals  $e_L$  and  $e_R$  to be outputted from line output ends LO(L) and LO(R), double talk detecting means 504 detects the double talk where sounds other than those reproduced by the

loudspeakers SP(L) and SP(R) are inputted into the  
microphones MC(L) and MC(R). The transfer function  
calculating means 502 makes relatively longer an update  
period of the filter characteristics of the filter means 40-1  
5 to 40-4 while the double talk is detected, whereas makes  
relatively shorter the update period of the filter  
characteristics while the double talk is not detected. This  
makes it possible to fully converge estimated errors when the  
double talk exists, and further, quicken convergence of  
10 estimated errors when there is no double talk.

[0146]

An orthogonalization process of the orthogonalizing  
filter 500 will be described. The orthogonalization process  
is implemented per prescribed time period of the input stereo  
15 signals. Herein, the orthogonalization process is carried  
out per frame (e.g. 512 samples) as shown in Fig. 29 (however,  
overlap processing may be implemented as shown in Fig. 49  
which will be described later). A vector having as its  
elements a sample group of one frame of the left-channel  
20 input signal  $x_L$  inputted into the orthogonalizing filter 500,  
and a vector having as its elements a sample group of one  
frame of the right-channel input signal  $x_R$  are given as  
[Equation 1]

$$\begin{aligned}\overrightarrow{x_L} &= (x_{L1}, x_{L2}, x_{L3}, \dots, x_{Ln}) \\ \overrightarrow{x_R} &= (x_{R1}, x_{R2}, x_{R3}, \dots, x_{Rn})\end{aligned}$$

Since  $x_L$  and  $x_R$  are the stereo signals, they mutually have a correlation. The orthogonalization process is carried out by performing a principal component analysis, using  $x_L$  and  $x_R$  as two variables, with respect to a sample group  
5 composed of combinations of such two variables per frame so as to derive eigenvectors of the first principal component and the second principal component that are mutually orthogonal, and projecting the respective samples composed of combinations of such two variables onto the derived  
10 eigenvectors of the first principal component and the second principal component, respectively.

[0147]

Concrete contents of calculation of the orthogonalization process will be described. Now, assuming  
15 that an observation matrix  $B$  is given as  
[Equation 2]

$$B = \begin{pmatrix} x_{L1}, x_{L2}, x_{L3}, \dots, x_{Ln} \\ x_{R1}, x_{R2}, x_{R3}, \dots, x_{Rn} \end{pmatrix}$$

20 then, a covariance matrix  $S$  of  $B$  becomes  
[Equation 3]

$$S = \frac{1}{n-1} BB^T \quad (B^T \text{ is a transposed matrix of } B)$$

$$= \frac{1}{n-1} \begin{pmatrix} x_{L1}, x_{L2}, x_{L3}, \dots, x_{Ln} \\ x_{R1}, x_{R2}, x_{R3}, \dots, x_{Rn} \end{pmatrix} \begin{pmatrix} x_{L1} & x_{R1} \\ x_{L2} & x_{R2} \\ \cdot & \cdot \\ \cdot & \cdot \\ x_{Ln} & x_{Rn} \end{pmatrix}$$

$$= \frac{1}{n-1} \begin{pmatrix} \sum_{k=1}^n x_{Lk}^2 & \sum_{k=1}^n x_{Lk} x_{Rk} \\ \sum_{k=1}^n x_{Lk} x_{Rk} & \sum_{k=1}^n x_{Rk}^2 \end{pmatrix}$$

$$= \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

( $S_{11}$  is variance of  $x_L$ ,  $S_{22}$  is variance of  $x_R$ , and  $S_{12}(=S_{21})$  is covariance of  $x_L$  and  $x_R$ )

5 Hence,

from

[Equation 4]

$$\begin{pmatrix} S_{11} - \lambda & S_{12} \\ S_{21} & S_{22} - \lambda \end{pmatrix} = 0$$

10

[Equation 5]

$$(S_{11} - \lambda)(S_{22} - \lambda) - S_{12}S_{21} = 0$$

is solved,

thus

15 [Equation 6]

$$\lambda = \frac{S_{11} + S_{22} \pm \sqrt{(S_{11} + S_{22})^2 - 4(S_{11}S_{22} - S_{12}^2)}}{2}$$

so that two solutions for eigenvalues  $\lambda$  are derived.

[0148]

Assuming that one of the two eigenvalues having greater variance (eigenvalue of the first principal component) is  $\lambda_1$ , an eigenvector  $U_{\max}$  corresponding to the

[Equation 8]

$$\begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix}$$

which establishes

[Equation 7]

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = \lambda_1 \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix}$$

where  $u_{11}^2 + u_{12}^2 = 1$

By solving  $u_{11}$  and  $u_{12}$ ,

[Equation 9]

$$u_{11} = \pm \frac{S_{12}}{\sqrt{S_{12}^2 - (\lambda_1 - S_{11})^2}}$$

$$u_{12} = \pm \frac{\lambda_1 - S_{11}}{\sqrt{S_{12}^2 - (\lambda_1 - S_{11})^2}}$$

(double signs in same order)

are derived. Regardless of the sign of  $u_{11}$  and  $u_{12}$  being plus or minus, the axis represented by the first principal

component is the same.

[0149]

On the other hand, assuming that one of the two eigenvalues having smaller variance (eigenvalue of the second principal component) is  $\lambda_2$ , an eigenvector  $U_{\min}$  corresponding to the eigenvalue  $\lambda_2$  is

[Equation 11]

$$\begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix}$$

10

which establishes

[Equation 10]

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = \lambda_2 \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix}$$

where  $u_{21}^2 + u_{22}^2 = 1$

15

By solving  $u_{21}$  and  $u_{22}$ ,

[Equation 12]

$$u_{21} = \pm \frac{S_{12}}{\sqrt{S_{12}^2 - (\lambda_2 - S_{11})^2}}$$
$$u_{22} = \pm \frac{\lambda_2 - S_{11}}{\sqrt{S_{12}^2 - (\lambda_2 - S_{11})^2}}$$

(double signs in same order)

20

are derived. Regardless of the sign of  $u_{21}$  and  $u_{22}$  being plus or minus, the axis represented by the second principal

component is the same.

From the foregoing, the eigenvector  $U_{\max}$  of the first principal component and the eigenvector  $U_{\min}$  of the second principal component are derived as follows.

5 [Equation 13]

$$\overrightarrow{U_{\max}} = \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix}, \quad \overrightarrow{U_{\min}} = \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix}$$

[0150]

10 With respect to the eigenvectors  $U_{\max}$  and  $U_{\min}$  obtained from the covariance matrix, it is not possible to predict in which quadrant it appears due to its nature. If the quadrant in which the eigenvector  $U_{\max}$  appears changes per frame,  $U_{\max}$ 's mutually cancel themselves upon ensemble-averaging  $U_{\max}$ 's per block later. If the quadrant in which the eigenvector  $U_{\min}$  appears changes per frame,  $U_{\min}$ 's mutually cancel themselves upon ensemble-averaging  $U_{\min}$ 's per block later. Therefore, a conversion operation is performed for fixing the quadrants in which the eigenvectors  $U_{\max}$  and  $U_{\min}$  appear. For example, when fixing the eigenvector  $U_{\max}$  to the first quadrant and the eigenvector  $U_{\min}$  to the fourth quadrant, it can be realized by the following conversion operation.

Of the eigenvectors  $U_{\max}$  and  $U_{\min}$ ,

25 • one existing in the first quadrant (positive, positive) or the third quadrant (negative, negative) is given as  $U_{\max}'$ , and

• one existing in the second quadrant (negative,

positive) or the fourth quadrant (positive, negative) is given as  $U_{\min}'$ . Then, conversion is executed such that

- when  $U_{\max}'$  exists in the first quadrant,  $U_{\max}=U_{\max}'$
- when  $U_{\max}'$  exists in the third quadrant,  $U_{\max}=-U_{\max}'$
- 5      • when  $U_{\min}'$  exists in the second quadrant,  $U_{\min}=-U_{\min}'$
- when  $U_{\min}'$  exists in the fourth quadrant,  $U_{\min}=U_{\min}'$ .

Through such a conversion operation, the quadrants in which the eigenvectors  $U_{\max}$  and  $U_{\min}$  appear can be fixed.

[0151]

10      Onto the eigenvector  $U_{\max}$  of the first principal component and the eigenvector  $U_{\min}$  of the second principal component that are obtained as described above, a column vector of the observation matrix B given as

[Equation 14]

15

$$\vec{b} = \begin{pmatrix} x_{Ln} \\ x_{Rn} \end{pmatrix}$$

is projected. A value of an output signal  $x_{\max}$  obtained by projecting the observation matrix B onto the eigenvector  $U_{\max}$

20      is derived as

[Equation 15]

$$x_{\max} = \vec{b} \bullet \overrightarrow{U_{\max}} \quad (\bullet \text{ represents inner product})$$

25      Further, a value of an output signal  $x_{\min}$  obtained by projecting the observation matrix B onto the eigenvector  $U_{\min}$  is derived as

[Equation 16]

$$x_{\min} = \vec{b} \bullet \overrightarrow{U_{\min}} \quad (\bullet \text{ represents inner product})$$

5

[0152]

The transfer function calculation process of the transfer function calculating means 502 will be described. (In case of Fixed Type Operation)

10 A filter characteristic of the orthogonalizing filter 500 is given as U. The filter characteristic U is a characteristic that projects the input signals  $x_L$  and  $x_R$  onto the mutually uncorrelated signals  $x_{\max}$  and  $x_{\min}$  based on the principal component analysis, and the following relationship is established.

15

[Equation 17]

$$\left. \begin{array}{l} \left( \begin{array}{c} \overrightarrow{x_{\max}} \\ \overrightarrow{x_{\min}} \end{array} \right) = U \left( \begin{array}{c} \overrightarrow{x_L} \\ \overrightarrow{x_R} \end{array} \right) \\ U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \end{array} \right\} \dots (95)$$

20 An inverse characteristic of the filter characteristic U is given as V. The inverse filter characteristic V is a characteristic that restores the signals  $x_{\max}$  and  $x_{\min}$  to the original signals  $x_L$  and  $x_R$ , and the following relationship is established.

[Equation 18]

25

$$\left. \begin{aligned} UV=VU=I \text{ (I is a unit matrix)} \\ V=\begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \end{aligned} \right\} \dots (96)$$

[0153]

Output signals  $y_L$  and  $y_R$  of the microphones MC(L) and  
5 MC(R) are expressed as

$$Y_L = H_{LL} \cdot X_L + H_{RL} \cdot X_R \dots (97)$$

$$Y_R = H_{LR} \cdot X_L + H_{RR} \cdot X_R \dots (98) .$$

From the equations (95) and (96),

[Equation 19]

10

$$\begin{pmatrix} x_L \\ x_R \end{pmatrix} = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{pmatrix} x_{\max} \\ x_{\min} \end{pmatrix} \dots (99)$$

hence, assuming that frequency-axis expressions of  $x_{\max}$  and  
 $x_{\min}$  are respectively given as  $X_{\max}$  and  $X_{\min}$ , the equations (97)

15 and (98) respectively become

$$Y_L = H_{LL}(v_{11} \cdot X_{\max} + v_{12} \cdot X_{\min}) + H_{RL}(v_{21} \cdot X_{\max} + v_{22} \cdot X_{\min}) \dots (97')$$

$$Y_R = H_{LR}(v_{11} \cdot X_{\max} + v_{12} \cdot X_{\min}) + H_{RR}(v_{21} \cdot X_{\max} + v_{22} \cdot X_{\min}) \dots (98') .$$

[0154]

When both sides of the equation (97') are multiplied  
20 by complex conjugates  $X_{\max}^*$  and  $X_{\min}^*$  of  $X_{\max}$  and  $X_{\min}$  (i.e.  
deriving cross spectra) and ensemble-averaged, because  $X_{\max}$   
and  $X_{\min}$  are mutually orthogonal, expected values of  $X_{\max}^* \cdot X_{\min}$   
and  $X_{\min}^* \cdot X_{\max}$  become zero, respectively, so that the  
following two equations are obtained (note:  $E[ ]$  represents  
25 the ensemble average).

$$E[X_{\max}^* \cdot Y_L] = E[X_{\max}^* \cdot H_{LL} \cdot v_{11} \cdot X_{\max} + X_{\max}^* \cdot H_{RL} \cdot v_{21} \cdot X_{\max}] \dots (100)$$

$$E[X_{\min}^* \cdot Y_L] = E[X_{\min}^* \cdot H_{LL} \cdot v_{12} \cdot X_{\min} + X_{\min}^* \cdot H_{RL} \cdot v_{22} \cdot X_{\min}] \dots (101)$$

5 Similarly, when both sides of the equation (98') are multiplied by complex conjugates  $X_{\max}^*$  and  $X_{\min}^*$  of  $X_{\max}$  and  $X_{\min}$  and ensemble-averaged, the following two equations are obtained.

$$10 \quad E[X_{\max}^* \cdot Y_R] = E[X_{\max}^* \cdot H_{LR} \cdot v_{11} \cdot X_{\max} + X_{\max}^* \cdot H_{RR} \cdot v_{21} \cdot X_{\max}] \dots (102)$$

$$E[X_{\min}^* \cdot Y_R] = E[X_{\min}^* \cdot H_{LR} \cdot v_{12} \cdot X_{\min} + X_{\min}^* \cdot H_{RR} \cdot v_{22} \cdot X_{\min}] \dots (103)$$

[0155]

15 Here, if a change of the eigenvalues  $\lambda$  is small in a time period of performing ensemble averaging,  
[Equation 20]

$$\left. \begin{aligned} U &= \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \approx E[U] = \begin{pmatrix} E[u_{11}] & E[u_{12}] \\ E[u_{21}] & E[u_{22}] \end{pmatrix} \\ V &= \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \approx E[V] = \begin{pmatrix} E[v_{11}] & E[v_{12}] \\ E[v_{21}] & E[v_{22}] \end{pmatrix} \end{aligned} \right\} \dots (104)$$

20 is established, so that the equations (100) to (103) are rewritten as the following equations (100') to (103').

$$E[X_{\max}^* \cdot Y_L] \approx E[|X_{\max}|^2] \cdot H_{LL} \cdot E[v_{11}] + E[|X_{\max}|^2] \cdot H_{RL} \cdot E[v_{21}] \dots (100')$$

$$E[X_{\min}^* \cdot Y_L] \approx E[|X_{\min}|^2] \cdot H_{LL} \cdot E[v_{12}] + E[|X_{\min}|^2] \cdot H_{RL} \cdot E[v_{22}] \dots (101')$$

25

$$\begin{aligned} E[X^*_{\max} \cdot Y_R] &\approx E[|X_{\max}|^2] \cdot H_{LR} \cdot E[v_{11}] \\ &\quad + E[|X_{\max}|^2] \cdot H_{RR} \cdot E[v_{21}] \dots (102') \end{aligned}$$

$$\begin{aligned} E[X^*_{\min} \cdot Y_R] &\approx E[|X_{\min}|^2] \cdot H_{LR} \cdot E[v_{12}] \\ &\quad + E[|X_{\min}|^2] \cdot H_{RR} \cdot E[v_{22}] \dots (103') \end{aligned}$$

5 [0156]

When both sides of the equations (100') and (102') are divided by  $E[|X_{\max}|^2]$  and both sides of the equations (101') and (103') are divided by  $E[|X_{\min}|^2]$ , respectively,

$$\begin{aligned} 10 \quad E[X^*_{\max} \cdot Y_L] / E[|X_{\max}|^2] &\approx H_{LL} \cdot E[v_{11}] + H_{RL} \cdot E[v_{21}] \\ E[X^*_{\min} \cdot Y_L] / E[|X_{\min}|^2] &\approx H_{LL} \cdot E[v_{12}] + H_{RL} \cdot E[v_{22}] \\ E[X^*_{\max} \cdot Y_R] / E[|X_{\max}|^2] &\approx H_{LR} \cdot E[v_{11}] + H_{RR} \cdot E[v_{21}] \\ E[X^*_{\min} \cdot Y_R] / E[|X_{\min}|^2] &\approx H_{LR} \cdot E[v_{12}] + H_{RR} \cdot E[v_{22}] \end{aligned}$$

hence

[Equation 21]

$$15 \quad \begin{pmatrix} \frac{E[X^*_{\max} \cdot Y_L]}{E[|X_{\max}|^2]} & \frac{E[X^*_{\min} \cdot Y_L]}{E[|X_{\min}|^2]} \\ \frac{E[X^*_{\max} \cdot Y_R]}{E[|X_{\max}|^2]} & \frac{E[X^*_{\min} \cdot Y_R]}{E[|X_{\min}|^2]} \end{pmatrix} \approx \begin{pmatrix} H_{LL} & H_{RL} \\ H_{LR} & H_{RR} \end{pmatrix} \begin{pmatrix} E[v_{11}] & E[v_{12}] \\ E[v_{21}] & E[v_{22}] \end{pmatrix} \dots (105)$$

is obtained. From  $E[U] \cdot E[V] \approx I$ , the equation (105) is rewritten as

20 [Equation 22]

$$\begin{pmatrix} H_{LL} & H_{RL} \\ H_{LR} & H_{RR} \end{pmatrix} \approx \begin{pmatrix} \frac{E[X^*_{\max} \cdot Y_L]}{E[|X_{\max}|^2]} & \frac{E[X^*_{\min} \cdot Y_L]}{E[|X_{\min}|^2]} \\ \frac{E[X^*_{\max} \cdot Y_R]}{E[|X_{\max}|^2]} & \frac{E[X^*_{\min} \cdot Y_R]}{E[|X_{\min}|^2]} \end{pmatrix} \begin{pmatrix} E[u_{11}] & E[u_{12}] \\ E[u_{21}] & E[u_{22}] \end{pmatrix} \dots (106)$$

[0157]

Impulse responses  $h_{LL}$ ,  $h_{RL}$ ,  $h_{LR}$  and  $h_{RR}$  obtained by applying the inverse Fourier transformation to the transfer functions  $H_{LL}$ ,  $H_{RL}$ ,  $H_{LR}$  and  $H_{RR}$  derived from the equation (106) are the filter characteristics to be set to the filter means 40-1, 40-2, 40-3 and 40-4, respectively. Therefore, the transfer function calculating means 502 derives the respective transfer functions  $H_{LL}$ ,  $H_{RL}$ ,  $H_{LR}$  and  $H_{RR}$  based on the signals  $x_{\max}$  and  $x_{\min}$  outputted from the orthogonalizing filter 500, the filter characteristic U of the orthogonalizing filter 500 and the output signals  $y_L$  and  $y_R$  of the microphones MC(L) and MC(R), derives the impulse responses  $h_{LL}$ ,  $h_{RL}$ ,  $h_{LR}$  and  $h_{RR}$  by applying the inverse Fourier transformation to those derived transfer functions, sets the derived impulse responses to the filter means 40-1, 40-2, 40-3 and 40-4, respectively, and further, updates the impulse responses by repeating this calculation per suitably determined prescribed time period (e.g. time period of performing ensemble averaging).

20 [0158]

(In case of Adaptive Type Operation)

Assuming that the filter characteristics set to the filter means 40-1, 40-2, 40-3 and 40-4 are given as  $\hat{H}_{LL}$ ,  $\hat{H}_{RL}$ ,  $\hat{H}_{LR}$  and  $\hat{H}_{RR}$  ( $\hat{h}_{LL}$ ,  $\hat{h}_{RL}$ ,  $\hat{h}_{LR}$  and  $\hat{h}_{RR}$  when expressed in terms of the impulse responses), the signals  $e_L$  and  $e_R$  outputted from the subtractors 48 and 50 of Fig. 28 are expressed as

$$E_L = (H_{LL} \cdot X_L + H_{RL} \cdot X_R) - (\hat{H}_{LL} \cdot X_L + \hat{H}_{RL} \cdot X_R) \dots (107)$$

$$E_R = (H_{LR} \cdot X_L + H_{RR} \cdot X_R) - (H_{LR}^{\wedge} \cdot X_L + H_{RR}^{\wedge} \cdot X_R) \dots (108).$$

From the foregoing equation (99), the equations (107) and (108) respectively become

$$E_L = (H_{LL} - H_{LL}^{\wedge}) (v_{11} \cdot X_{\max} + v_{12} \cdot X_{\min}) + (H_{RL} - H_{RL}^{\wedge}) (v_{21} \cdot X_{\max} + v_{22} \cdot X_{\min}) \dots (107')$$

$$E_R = (H_{LR} - H_{LR}^{\wedge}) (v_{11} \cdot X_{\max} + v_{12} \cdot X_{\min}) + (H_{RR} - H_{RR}^{\wedge}) (v_{21} \cdot X_{\max} + v_{22} \cdot X_{\min}) \dots (108').$$

[0159]

When the estimated errors of the transfer functions are given as

$$\Delta H_{LL} = H_{LL} - H_{LL}^{\wedge}$$

$$\Delta H_{RL} = H_{RL} - H_{RL}^{\wedge}$$

$$\Delta H_{LR} = H_{LR} - H_{LR}^{\wedge}$$

$$\Delta H_{RR} = H_{RR} - H_{RR}^{\wedge}$$

the equations (107') and (108') respectively become

$$E_L = \Delta H_{LL} (v_{11} \cdot X_{\max} + v_{12} \cdot X_{\min}) + \Delta H_{RL} (v_{21} \cdot X_{\max} + v_{22} \cdot X_{\min}) \dots (107'')$$

$$E_R = \Delta H_{LR} (v_{11} \cdot X_{\max} + v_{12} \cdot X_{\min}) + \Delta H_{RR} (v_{21} \cdot X_{\max} + v_{22} \cdot X_{\min}) \dots (108'').$$

[0160]

When both sides of the equation (107'') are multiplied by complex conjugates  $X_{\max}^*$  and  $X_{\min}^*$  of  $X_{\max}$  and  $X_{\min}$  (i.e. deriving cross spectra) and ensemble-averaged, because  $X_{\max}$  and  $X_{\min}$  are mutually orthogonal, expected values of  $X_{\max}^* \cdot X_{\min}$  and  $X_{\min}^* \cdot X_{\max}$  become zero, respectively, so that the following two equations are obtained (note:  $E[ ]$  represents the ensemble average).

$$E[X_{\max}^* \cdot E_L] = E[X_{\max}^* \cdot \Delta H_{LL} \cdot v_{11} \cdot X_{\max} + X_{\max}^* \cdot \Delta H_{RL} \cdot v_{21} \cdot X_{\max}] \dots (109)$$

$$E[X_{\min}^* \cdot E_L] = E[X_{\min}^* \cdot \Delta H_{LL} \cdot v_{12} \cdot X_{\min} + X_{\min}^* \cdot \Delta H_{RL} \cdot v_{22} \cdot X_{\min}] \dots (110)$$

5 Similarly, when both sides of the equation (108") are multiplied by complex conjugates  $X_{\max}^*$  and  $X_{\min}^*$  of  $X_{\max}$  and  $X_{\min}$  and ensemble-averaged, the following two equations are obtained.

$$10 \quad E[X_{\max}^* \cdot E_R] = E[X_{\max}^* \cdot \Delta H_{LR} \cdot v_{11} \cdot X_{\max} + X_{\max}^* \cdot \Delta H_{RR} \cdot v_{21} \cdot X_{\max}] \dots (111)$$

$$E[X_{\min}^* \cdot E_R] = E[X_{\min}^* \cdot \Delta H_{LR} \cdot v_{12} \cdot X_{\min} + X_{\min}^* \cdot \Delta H_{RR} \cdot v_{22} \cdot X_{\min}] \dots (112)$$

[0161]

15 Here, if a change of the eigenvalues  $\lambda$  is small in a time period of performing ensemble averaging, the foregoing equation (104) is established so that the equations (109) to (112) are rewritten as the following equations (109') to (112').

$$20 \quad E[X_{\max}^* \cdot E_L] \approx E[|X_{\max}|^2] \cdot \Delta H_{LL} \cdot E[v_{11}] + E[|X_{\max}|^2] \cdot \Delta H_{RL} \cdot E[v_{21}] \dots (109')$$

$$E[X_{\min}^* \cdot E_L] \approx E[|X_{\min}|^2] \cdot \Delta H_{LL} \cdot E[v_{12}] + E[|X_{\min}|^2] \cdot \Delta H_{RL} \cdot E[v_{22}] \dots (110')$$

$$E[X_{\max}^* \cdot E_R] \approx E[|X_{\max}|^2] \cdot \Delta H_{LR} \cdot E[v_{11}] + E[|X_{\max}|^2] \cdot \Delta H_{RR} \cdot E[v_{21}] \dots (111')$$

$$25 \quad E[X_{\min}^* \cdot E_R] \approx E[|X_{\min}|^2] \cdot \Delta H_{LR} \cdot E[v_{12}] + E[|X_{\min}|^2] \cdot \Delta H_{RR} \cdot E[v_{22}] \dots (112')$$

[0162]

When both sides of the equations (109') and (111') are divided by  $E[|X_{\max}|^2]$  and both sides of the equations (110') and (112') are divided by  $E[|X_{\min}|^2]$ , respectively,

$$\begin{aligned} E[X^*_{\max} \cdot E_L] / E[|X_{\max}|^2] &\approx \Delta H_{LL} \cdot E[v_{11}] + \Delta H_{RL} \cdot E[v_{21}] \\ E[X^*_{\min} \cdot E_L] / E[|X_{\min}|^2] &\approx \Delta H_{LL} \cdot E[v_{12}] + \Delta H_{RL} \cdot E[v_{22}] \\ E[X^*_{\max} \cdot E_R] / E[|X_{\max}|^2] &\approx \Delta H_{LR} \cdot E[v_{11}] + \Delta H_{RR} \cdot E[v_{21}] \\ E[X^*_{\min} \cdot E_R] / E[|X_{\min}|^2] &\approx \Delta H_{LR} \cdot E[v_{12}] + \Delta H_{RR} \cdot E[v_{22}] \end{aligned}$$

hence

[Equation 23]

10

$$\begin{pmatrix} \frac{E[X^*_{\max} \cdot E_L]}{E[|X_{\max}|^2]} & \frac{E[X^*_{\min} \cdot E_L]}{E[|X_{\min}|^2]} \\ \frac{E[X^*_{\max} \cdot E_R]}{E[|X_{\max}|^2]} & \frac{E[X^*_{\min} \cdot E_R]}{E[|X_{\min}|^2]} \end{pmatrix} \approx \begin{pmatrix} \Delta H_{LL} & \Delta H_{RL} \\ \Delta H_{LR} & \Delta H_{RR} \end{pmatrix} \begin{pmatrix} E[v_{11}] & E[v_{12}] \\ E[v_{21}] & E[v_{22}] \end{pmatrix} \dots (113)$$

is obtained. From  $E[U] \cdot E[V] \approx I$ , the equation (113) is rewritten as

15

[Equation 24]

$$\begin{pmatrix} \Delta H_{LL} & \Delta H_{RL} \\ \Delta H_{LR} & \Delta H_{RR} \end{pmatrix} \approx \begin{pmatrix} \frac{E[X^*_{\max} \cdot E_L]}{E[|X_{\max}|^2]} & \frac{E[X^*_{\min} \cdot E_L]}{E[|X_{\min}|^2]} \\ \frac{E[X^*_{\max} \cdot E_R]}{E[|X_{\max}|^2]} & \frac{E[X^*_{\min} \cdot E_R]}{E[|X_{\min}|^2]} \end{pmatrix} \begin{pmatrix} E[u_{11}] & E[u_{12}] \\ E[u_{21}] & E[u_{22}] \end{pmatrix} \dots (114)$$

[0163]

20

Using the estimated errors  $\Delta H_{LL}$ ,  $\Delta H_{RL}$ ,  $\Delta H_{LR}$  and  $\Delta H_{RR}$  derived from the equation (114), the filter characteristics of the filter means 40-1, 40-2, 40-3 and 40-4 are updated per suitably determined prescribed time period (e.g. time period

of performing ensemble averaging). For example, assuming that impulse responses  $h_{LL}$ ,  $h_{RL}$ ,  $h_{LR}$  and  $h_{RR}$  after  $K$ -th updating are given as  $h_{LL}(k)$ ,  $h_{RL}(k)$ ,  $h_{LR}(k)$  and  $h_{RR}(k)$ , using impulse responses  $\Delta h_{LL}$ ,  $\Delta h_{RL}$ ,  $\Delta h_{LR}$  and  $\Delta h_{RR}$  corresponding to  
5 the derived estimated errors  $\Delta H_{LL}$ ,  $\Delta H_{RL}$ ,  $\Delta H_{LR}$  and  $\Delta H_{RR}$ ,

$$h_{LL}(k+1) = h_{LL}(k) + \alpha \Delta h_{LL}$$

$$h_{RL}(k+1) = h_{RL}(k) + \alpha \Delta h_{RL}$$

$$h_{LR}(k+1) = h_{LR}(k) + \alpha \Delta h_{LR}$$

$$h_{RR}(k+1) = h_{RR}(k) + \alpha \Delta h_{RR}.$$

10 Using these updating equations,  $(k+1)$ th impulse responses  $h_{LL}(k+1)$ ,  $h_{RL}(k+1)$ ,  $h_{LR}(k+1)$  and  $h_{RR}(k+1)$  are derived and set to the filter means 40-1, 40-2, 40-3 and 40-4, respectively, which is repeated per suitably determined prescribed time period (e.g. time period of performing  
15 ensemble averaging).

[0164]

Fig. 30 shows functional blocks of the orthogonalizing filter 500 of Fig. 28. The input stereo signals  $x_L$  and  $x_R$  are inputted from input ends 506 and 508,  
20 respectively. Covariance matrix calculating means 510 derives a covariance matrix  $S$  of the input stereo signals  $x_L$  and  $x_R$  per frame. Based on the derived covariance matrix  $S$ , eigenvector calculating means 512 derives coefficients  $u_{11}$ ,  $u_{12}$ ,  $u_{21}$  and  $u_{22}$  of eigenvectors  $U_{\max}$  and  $U_{\min}$  of the first and  
25 second principal components per frame. The derived coefficients  $u_{11}$ ,  $u_{21}$ ,  $u_{12}$  and  $u_{22}$  are set to coefficient multipliers 514, 516, 518 and 520, respectively. The

coefficient multipliers 514 and 516 give the coefficients  $u_{11}$  and  $u_{21}$  to the input signal  $x_L$  to derive  $x_L \cdot u_{11}$  and  $x_L \cdot u_{21}$ , respectively. The coefficient multipliers 518 and 520 give the coefficients  $u_{12}$  and  $u_{22}$  to the input signal  $x_R$  to derive  $x_R \cdot u_{12}$  and  $x_R \cdot u_{22}$ , respectively. An adder 522 derives  $x_L \cdot u_{11} + x_R \cdot u_{12}$  as a signal  $x_{\max}$  obtained by projecting the input stereo signals  $x_L$  and  $x_R$  onto the eigenvector  $U_{\max}$ . An adder 524 derives  $x_L \cdot u_{21} + x_R \cdot u_{22}$  as a signal  $x_{\min}$  obtained by projecting the input stereo signals  $x_L$  and  $x_R$  onto the eigenvector  $U_{\min}$ . The signals  $x_{\max}$  and  $x_{\min}$  are respectively outputted from output ends 526 and 528, while the coefficients  $u_{11}$ ,  $u_{21}$ ,  $u_{12}$  and  $u_{22}$  of the eigenvectors are outputted from an output end 530.

[0165]

Fig. 31 shows functional blocks of the transfer function calculating means 502 of Fig. 28. The signals  $x_{\max}$  and  $x_{\min}$  are respectively inputted from input ends 532 and 534, while the coefficients  $u_{11}$ ,  $u_{21}$ ,  $u_{12}$  and  $u_{22}$  of the eigenvectors are inputted from an input end 536. FFT means 538 applies the fast Fourier transformation to the signal  $x_{\max}$ . Complex conjugate calculating means 540 calculates a complex conjugate  $X_{\max}^*$  of  $X_{\max}$ . Power spectrum calculating means 542 calculates  $X_{\max} \cdot X_{\max}^* = |X_{\max}|^2$ . Ensemble averaging means 544 calculates  $E[|X_{\max}|^2]$ . FFT means 546 applies the fast Fourier transformation to the signal  $x_{\min}$ . Complex conjugate calculating means 548 calculates a complex conjugate  $X_{\min}^*$  of  $X_{\min}$ . Power spectrum calculating means 550 calculates

$X_{\min} \cdot X_{\min}^* = |X_{\min}|^2$ . Ensemble averaging means 552 calculates  $E[|X_{\min}|^2]$ .

[0166]

An input end 554 is inputted with an output signal  $y_L$  of the microphone MC(L) (or output signal  $E_L$  of the subtracter 48). FFT means 558 applies the fast Fourier transformation to the signal  $y_L$  (or signal  $E_L$ ). Cross spectrum calculating means 560 derives  $X_{\max}^* \cdot Y_L$  (or  $X_{\max}^* \cdot E_L$ ), and ensemble averaging means 562 derives  $E[X_{\max}^* \cdot Y_L]$  (or  $E[X_{\max}^* \cdot E_L]$ ). Cross spectrum calculating means 564 derives  $X_{\min}^* \cdot Y_L$  (or  $X_{\min}^* \cdot E_L$ ), and ensemble averaging means 566 derives  $E[X_{\min}^* \cdot Y_L]$  (or  $E[X_{\min}^* \cdot E_L]$ ).

[0167]

An input end 568 is inputted with an output signal  $y_R$  of the microphone MC(R) (or output signal  $E_R$  of the subtracter 50). FFT means 570 applies the fast Fourier transformation to the signal  $y_R$  (or signal  $E_R$ ). Cross spectrum calculating means 572 derives  $X_{\max}^* \cdot Y_R$  (or  $X_{\max}^* \cdot E_R$ ), and ensemble averaging means 574 derives  $E[X_{\max}^* \cdot Y_R]$  (or  $E[X_{\max}^* \cdot E_R]$ ). Cross spectrum calculating means 576 derives  $X_{\min}^* \cdot Y_R$  (or  $X_{\min}^* \cdot E_R$ ), and ensemble averaging means 578 derives  $E[X_{\min}^* \cdot Y_R]$  (or  $E[X_{\min}^* \cdot E_R]$ ).

[0168]

Composite transfer function calculating means 580 derives the first term on the right side of the equation (106) {or the equation (114)} based on  $E[|X_{\max}|^2]$ ,  $E[|X_{\min}|^2]$ ,  $E[X_{\max}^* \cdot Y_L]$  (or  $E[X_{\max}^* \cdot E_L]$ ),  $E[X_{\min}^* \cdot Y_L]$  (or  $E[X_{\min}^* \cdot E_L]$ ),

$E[X^*_{\max} \cdot Y_R]$  (or  $E[X^*_{\max} \cdot E_R]$ ), and  $E[X^*_{\min} \cdot Y_R]$  (or  $E[X^*_{\min} \cdot E_R]$ ) which are derived as described above. Averaging means 582 averages the coefficients  $u_{11}$ ,  $u_{21}$ ,  $u_{12}$  and  $u_{22}$  of the eigenvectors individually to derive  $E[v_{11}]$ ,  $E[v_{12}]$ ,  $E[v_{21}]$  and  $E[v_{22}]$ .

5 [0169]

Based on the outputs of the composite transfer function calculating means 580 and the averaging means 582, transfer function calculating means 584 performs a calculation of the right side of the equation (106) {or the  
10 equation (114)} to derive individual transfer functions  $H_{LL}$ ,  $H_{RL}$ ,  $H_{LR}$  and  $H_{RR}$  (or their estimated errors  $\Delta H_{LL}$ ,  $\Delta H_{RL}$ ,  $\Delta H_{LR}$  and  $\Delta H_{RR}$ ). Inverse FFT means 586 applies the inverse fast Fourier transformation to the derived transfer functions  $H_{LL}$ ,  $H_{RL}$ ,  $H_{LR}$  and  $H_{RR}$  (or their estimated errors  $\Delta H_{LL}$ ,  $\Delta H_{RL}$ ,  $\Delta H_{LR}$  and  
15  $\Delta H_{RR}$ ) to derive corresponding impulse responses  $h_{LL}$ ,  $h_{RL}$ ,  $h_{LR}$  and  $h_{RR}$  (or their estimated errors  $\Delta h_{LL}$ ,  $\Delta h_{RL}$ ,  $\Delta h_{LR}$  and  $\Delta h_{RR}$ ), and outputs them from output ends 588, 590, 592 and 594, respectively. As shown in Fig. 32, the transfer function calculating means 584 and the inverse FFT means 586 may be  
20 exchanged therebetween in their positions.

[0170]

With respect to the thus structured stereo echo canceller 16, 24 of Fig. 28, the results of carrying out adaptive type operation simulations are shown in Figs. 33 to  
25 48 in connection with one audio transfer system. Figs. 33 to 40 show time-domain variations in echo cancellation amount, while Figs. 41 to 48 show time-domain variations in transfer

function estimated error. Herein, the simulations were conducted under the following conditions.

- Sampling Frequency: 11.025kHz

- The Number of Samples in One Frame: 4096 samples

5     • The Number of Frames in One Block: variable (2 frames, 4 frames, 8 frames, 16 frames)

10    • Update Period of Filter Characteristic: per block (about 0.75seconds in case of the number of frames in one block being two, about 1.5 seconds in case of 4 frames, about 3 seconds in case of 8 frames, about 6 seconds in case of 16 frames)

15    • The Mean Number of Times of Ensemble Averaging in One Block: 31 times in case of the number of frames in one block being two, 63 times in case of 4 frames, 127 times in case of 8 frames, 255 times in case of 16 frames {For the purpose of increasing the mean number of times, as shown in Fig. 49, one frame is divided into 16 intervals, and data corresponding to one frame is extracted from the head of each divisional interval while overlapping data extraction successively, 20 thereby to derive individual parameter values to be ensemble-averaged, and the derived individual parameter values are respectively ensemble-averaged in one block. Therefore, the mean number of times N of ensemble averaging becomes such that  $N = (16 \times \text{the number of frames in one block} - 1).$ }

25           In any case, the filter characteristics are not set in the first block, and the initial setting is implemented in the second block, and thereafter, updating is performed per

block. The axis of ordinates (dB) is defined such that 0dB represents the initial state where the filter characteristics are not set. Differences in simulation condition with respect to Figs. 33 to 48 are as follows.

5 [Table 2]

Figure Number	Presence/Absence of Double Talk	Number of Frames in One Block
Fig. 33, Fig. 41	Absent	2
Fig. 34, Fig. 42		4
Fig. 35, Fig. 43		8
Fig. 36, Fig. 44		16
Fig. 37, Fig. 45	Present	2
Fig. 38, Fig. 46		4
Fig. 39, Fig. 47		8
Fig. 40, Fig. 48		16

[0171]

10 The simulation results of Figs. 33 to 48 are considered.

(1) In Case of Absence of Double Talk

15 Since the rising speed of the echo cancellation amount is Fig. 33 > Fig. 34 > Fig. 35 > Fig. 36, the rising speed of the echo cancellation amount becomes faster as one block (update period) becomes shorter (the number of frames per block becomes smaller). When the number of frames in one block is 2, 4 or 8, the echo cancellation amount of about

25dB is obtained (Fig. 33, 34 or 35). When the number of frames in one block increases, i.e. 16 frames, the echo cancellation amount requires a long time for reduction thereof (Fig. 36). Since the estimated error convergence speed is Fig. 41 > Fig. 42 > Fig. 43 > Fig. 44, the estimated error convergence speed becomes faster as one block becomes shorter.

(2) In Case of Presence of Double Talk

When the number of frames in one block is 2 or 4 frames, the echo cancellation amount is not increased (Fig. 37 or 38), and the estimated error is not converged (Fig. 45 or 46) and thus can not be estimated. When the number of frames in one block is 8, the echo cancellation amount of about 15dB is obtained (Fig. 39), and the estimated error is converged to about -6dB (Fig. 47). When the number of frames in one block is 16, the echo cancellation amount of about 17dB is obtained (Fig. 40), and the estimated error is converged to about 10dB and thus the fairly stable estimation can be achieved.

From the foregoing simulation results, the followings can be said.

(a) When the double talk is not detected, the convergence of the estimated error can be quickened by relatively shortening the update period of the filter characteristic.

(b) When the double talk is detected, the estimated error can be fully converged by relatively prolonging the

update period of the filter characteristic.

Therefore, as described above, the transfer function calculating means 502 of Fig. 28 makes relatively longer the update period of the filter characteristics of the filter

5 means 40-1 to 40-4 while the double talk is detected, whereas makes relatively shorter the update period of the filter characteristics while the double talk is not detected. This makes it possible to fully converge the estimated errors when the double talk exists, and further, quicken the convergence  
10 of the estimated errors when there is no double talk.

[0172]

Fig. 50 shows a modification of the stereo echo canceller 16, 24 of Fig. 28. The same symbols are used with respect to those portions common to Fig. 28. In this  
15 modification, an orthogonalizing filter is disposed on signal lines of loudspeakers SP(L) and SP(R). An inverse filter 596 has an inverse characteristic of the orthogonalizing filter 500 {the inverse filter characteristic V of the foregoing equation (96)}, thereby to restore output signals  $x_{\max}$  and  $x_{\min}$   
20 of the orthogonalizing filter 500 to the original signal  $x_L$  and  $x_R$ , and feeds them to the loudspeakers SP(L) and SP(R).

In the foregoing embodiments, the number of the loudspeakers is two and the number of the microphones is two. However, it may also be configured that the number of the  
25 loudspeakers is two, while the number of the microphones is one. Fig. 51 shows a structural example as a result of modifying Fig. 1 in such a manner. The same symbols are used

with respect to those portions common to Fig. 1. Left/right two-channel stereo signals  $x_L$  and  $x_R$  transmitted from the spot on the counterpart side and inputted into line input ends LI(L) and LI(R) are outputted from sound output ends SO(L) and SO(R) as they are (i.e. not through sum/difference signal producing means 52), and reproduced at loudspeakers SP(L) and SP(R), respectively.

Filter means 40-1 is set with an impulse response corresponding to a transfer function  $H_L$  between the loudspeaker SP(L) and a microphone MC and performs, using such an impulse response, a convolution calculation of a signal  $x_L$  to be outputted from the sound output end SO(L), thereby producing an echo cancel signal EC1 corresponding to a signal  $y_L$  obtained such that the signal  $x_L$  outputted from the sound output end SO(L) is reproduced at the loudspeaker SP(L), collected by the microphone MC and inputted into a sound input end SI. Filter means 40-3 is set with an impulse response corresponding to a transfer function  $H_R$  between the loudspeaker SP(R) and the microphone MC and performs, using such an impulse response, a convolution calculation of a signal  $x_R$  to be outputted from the sound output end SO(R), thereby producing an echo cancel signal EC3 corresponding to a signal  $y_R$  obtained such that the signal  $x_R$  outputted from the sound output end SO(R) is reproduced at the loudspeaker SP(R), collected by the microphone MC and inputted into the sound input end SI. An adder 44 performs a calculation of  $EC1+EC3$ . A subtracter 48 subtracts an echo cancel signal

EC1+EC3 from a collected audio signal  $y$  ( $=y_L+y_R$ ) of the microphone MC inputted from the sound input end SI, thereby to perform echo cancellation. An echo-canceled signal  $e$  ( $=e_L+e_R$ ) is outputted from a line output end LO and  
5 transmitted toward the spot on the counterpart side.

[0173]

The sum/difference signal producing means 52 performs addition, using an adder 54, of the left/right two-channel stereo signals  $x_L$  and  $x_R$  inputted into the line input  
10 ends LI(L) and LI(R) so as to produce a sum signal  $x_M$  ( $=x_L+x_R$ ), while performs subtraction thereof using a subtracter 56 so as to produce a difference signal  $x_S$  ( $=x_L-x_R$  (or it may also be  $x_R-x_L$ )). Transfer function calculating means 58 implements a cross-spectrum calculation between the sum  
15 signal  $x_M$  and the difference signal  $x_S$  produced by the sum/difference signal producing means 52 and the signal  $e$  outputted from the subtracter 48 and, based on this cross-spectrum calculation, sets filter characteristics (impulse responses) of the filter means 40-1 and 40-3. Specifically,  
20 upon starting the system, the filter characteristics of the filter means 40-1 and 40-3 are not set, i.e. coefficients are all set to zero, so that the echo cancel signals EC1 and EC3 are zero, and thus the collected audio signal of the microphone MC itself is outputted from the subtracter 48.  
25 Therefore, at this time, the transfer function calculating means 58 performs the cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_S$  produced by the

sum/difference signal producing means 52 and the collected audio signal  $e$  of the microphones MC outputted from the subtracter 48 and, based on this cross-spectrum calculation, derives transfer functions of two audio transfer systems

5 between the loudspeakers  $SP(L)$  and  $SP(R)$  and the microphone MC, respectively, and implements initial setting of the filter characteristics of the filter means 40-1 and 40-3 to values corresponding to such transfer functions. After the initial setting, since the echo cancel signals are produced

10 by the filter means 40-1 and 40-3, the echo cancel error signal  $e$  corresponding to a difference signal between the collected audio signal of the microphone MC and the echo cancel signal  $EC1+EC3$  is outputted from the subtracter 48. Therefore, at this time, the transfer function calculating

15 means 58 performs the cross-spectrum calculation between the sum signal  $x_M$  and the difference signal  $x_s$  produced by the sum/difference signal producing means 52 and the echo cancel error signal  $e$  outputted from the subtracter 48 and, based on this cross-spectrum calculation, derives estimated errors of

20 the transfer functions of the two audio transfer systems between the loudspeakers  $SP(L)$  and  $SP(R)$  and the microphone MC, respectively, and updates the filter characteristics of the filter means 40-1 and 40-3 to values that cancel the estimated errors, respectively. By repeating this updating

25 operation per prescribed time period, the echo cancel error can be converged to a minimum value. Further, even if the transfer functions change due to movement of the microphone

positions or the like, the echo cancel error can be converged to a minimum value by sequentially updating the filter characteristics of the filter means 40-1 and 40-3 depending thereon.

5 [0174]

Correlation detecting means 60 detects a correlation between the sum signal  $x_M$  and the difference signal  $x_S$  based on a correlation value calculation or the like, and stops updating of the foregoing filter characteristics when the correlation value is no less than a prescribed value. When the correlation value becomes lower than the prescribed value, updating of the foregoing filter characteristics is restarted. Also in the embodiments other than Fig. 1, it can be configured that the number of the loudspeakers is two, while the number of the microphones is one.

[0175]

The stereo echo canceller 16, 24 shown in each of the foregoing embodiments can be formed by the dedicated hardware or can also be realized through software processing in a general computer. For example, the functions of the respective blocks as shown in Fig. 1 and so forth can be accomplished by a CPU (Central Processing Unit) and storing means such as a RAM or ROM constituting the computer. Namely, the CPU may be caused to function as the echo canceller according to a program stored in the storing means such as the ROM or RAM.

[0176]

In the foregoing embodiments, the description has been made about the case where the two-channel stereo signals are handled. However, the echo cancellation can also be implemented using the technique of this invention with  
5 respect to those signals of three channels or more having a correlation with each other.